NATIONAL OPEN UNIVERSITY OF NIGERIA

COURSE CODE :BHM 733

COURSE TITLE:
E-BUSINESS SECURITY

COURSE GUIDE

BHM 733
E-BUSINESS SECURITY

Course Writer Gerald C. Okereke
Eco Communications Inc.
Lagos Ikeja

Programme Leader Dr. O. J. Onwe
National Open University of Nigeria

Course Coordinator Israel-Cookey
National Open University of Nigeria

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
No. 5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja
Nigeria

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published by
National Open University of Nigeria

Printed 2009

ISBN: 978-058-439-0

CONTENTS **PAGE**

Course Aims

This course has been designed to highlight the low points of e-business, and equip participants with corresponding strategies to fill-in these low points. At the end of this course, all the dimensions of e-business; institutional, legislative, educational, technical and otherwise securing should be well understood to e-business professionals and managers. The perspective of the course is to encourage businesses to go electronic despite the security challenges, by proffering practical and applicable solutions to security concerns.

Course Objectives

A summary of the objectives of this course includes to:

- appropriately define e-business and differentiate it from other related terms
- explain the various types of security measures for e-commerce
- have a basic understanding of how to set up security measures
- understand the needs for the security of e-business
- understand the problems encountered in e-business transactions
- define security
- identify and differentiate the types of controls
- learn how to implement network security and management
- explain the phases and how to plan for business recovery
- understand how to appraise management of risks and mitigation costs
- understand how to implement controls generally
- learn how to carry out Information System Audit Planning
- answer the question of the scope and justification in applying copyright
- understand how to obtain and enforce copyright
- appreciate the concerns of network security administrators

- define Internet firewalls
- identify the basic features of digital and electronic signatures
- understand the applications of digital and electronic signatures
- identify the applications of biometrics in security management
- identify the characteristics associated with biometrics
- identify the risks involved in e-business
- learn the different types of defense strategies against frauds
- understand the International Convention and Articles against plastic card frauds
- use Australia as a case study to understand legislative responses in dealing with e-frauds

Study Units

There are fourteen study units in this course:

Module 1

Unit 1 What is E-Business
Unit 2 Introduction to E-Business Security
Unit 3 E-Business Security Challenges

Unit 4 Types of Information Security Controls
Unit 5 Network Security and Management

Module 2

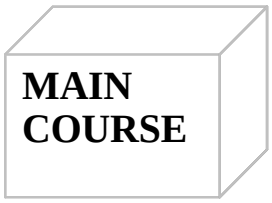
Unit 1 Business Continuity Planning
Unit 2 Information Systems Security Controls
Unit 3 Information Systems Audit
Unit 4 Copyright Law and Electronic Access to Information
Unit 5 Internet Firewall

Module 3

Unit 1 Digital Signature and Electronic Signature
Unit 2 Biometric Identification
Unit 3 Fraud Prevention
Unit 4 Sanctions against Plastic Card Fraud: The Case of Australia

Assessment

- The assignments represents 30% of the marks obtainable.
- Examination constitutes 70% of the marks obtainable.



Course Code BHM 733

Course Title E-Business Security

Course Writer Gerald C. Okereke
Eco Communications Inc.
Lagos Ikeja

Programme Leader Dr. O. J. Onwe
National Open University of Nigeria

Course Coordinator Israel-Cookey
National Open University of Nigeria

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
No. 5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja
Nigeria

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published by
National Open University of Nigeria

Printed 2009

ISBN: 978-058-439-0

CONTENTS PAGE

MODULE 1

Unit 1 What is E-Business?
Unit 2 Introduction to E-Business Security
Unit 3 Business Security Challenges
Unit 4 Types of Information Security Controls
Unit 5 Network Security and Management

UNIT 1 WHAT IS E-BUSINESS

CONTENTS

1.0 INTRODUCTION

Electronic Business, commonly referred to as "eBusiness" or "e-Business", may be defined as the utilisation of information and communication technologies (ICTs) in support of all the activities of business. Commerce constitutes the exchange of products and services between businesses, groups and individuals and hence can be seen as one of the essential activities of any business. Hence, electronic commerce or eCommerce focuses on the use of ICT to enable the external activities and relationships of the business with individuals, groups and other businesses.

Louis Gerstner, the former CEO of IBM, in his book, Who says *Elephants can't Dance attributes the term "e-Business" to IBM's* marketing and Internet teams in 1996.

Electronic business methods enable companies to link their internal and external data processing systems more efficiently and flexibly, to work more closely with suppliers and partners, and to better satisfy the needs and expectations of their customers.

In practice, e-business is more than just e-commerce. While e-business refers to more strategic focus with an emphasis on the functions that occur using electronic capabilities, e-commerce is a subset of an overall e-business strategy. E-commerce seeks to add revenue streams using the World Wide Web or the Internet to build and enhance relationships with clients and partners and to improve efficiency using the Empty Vessel strategy. Often, e-commerce involves the application of knowledge management systems.

E-business involves business processes spanning the entire value chain: electronic purchasing and supply chain management, processing orders electronically, handling customer service, and cooperating with business partners. Special technical standards for e-business facilitate the exchange of data between companies. E-business software solutions allow the integration of intra and inter firm business processes. E-business can be conducted using the Web, the Internet, intranets, extranets, or some combination of these.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- appropriately define e-business and differentiate it from other related terms
- differentiate e-commerce from e-business
- identify the various models of e-business
- explain the concepts and evolution of e-business.

3.0 MAIN CONTENT

3.1 Definitions

There are many definitions for e-business, e-commerce as well as related complementary terms. Some of these definitions are stated below. When people hear e-business, they generally think selling over the Web.

IBM which was one of the first users of the term defined e-business as business process, using web-based technology to help business streamline process, improve productivity, and increase efficiencies. It enables companies to easily communicate with partners, vendors and

customers, connect back-end data systems and transact commerce in a secure manner.

Electronic Business or e-business in short refers broadly to the use of technologies, particularly the Information and Communication Technologies (ICTs), to conduct or support to improve business activities and processes, including research and development, procurement, design and development, operation, manufacturing, marketing and sales, logistics, human resources management, finance, and value chain integration.

Perspective Definitions of E-Business

In details, e-business can be defined from the following perspectives:

• Communications: Delivery of goods, services, information, or payments over the computer networks or any other electronic means.

• Commercial (Trading): Provides capability of buying and selling products, services, and information on the Internet and via other online services.

• Business Process: Doing business electronically by completing business processes over electronic networks, thereby substituting information for physical business processes.

• Services: A tool that addresses the desire of governments, firms, consumers, and management to cut service costs while improving the quality of consumer service and increasing the speed of service delivery.

• Learning: An enabler of online training and education in schools, universities, and other organizations, including businesses.

• Collaborative: The framework for inter- and intra-organizational collaboration.

• Community: Provides a gathering place for community members to learn, transact, and collaborate.

Simply, e-business could be any system that suppliers, distributors, or customers use the ICT, particularly the Internet, as the basis for conducting their business operation, for example:

• Communicate with clients or suppliers via email;
• Send email to other organizations to order supplies;

- Sell or promote products or services via a web site and/or email;
- Publish a web site to provide public information about the business;
- Use the Internet for online banking and paying bills;
- Research information about customers and competitors using web sites;
- Provide technical or customer service by email or web site; and
- Manage and distribute internal organization documents via an intranet.

Also, E-Business is the conduction of business in the electronic marketplace. In practice, this involves the introduction of new revenue streams through the use of e-commerce, the enhancement of relationships between clients and partners and improving efficiency from using knowledge management systems. E-business can be conducted over the public Internet, through internal intranets and over secure private extranets.

- E-Commerce

E-Commerce is the term used to describe financial transactions in the electronic marketplace. This typically includes buying or selling products or services on the Internet, via secure networks or extranets. Transactions may be business-to-business or business to consumer.

- B2B

IDC defines B2B eCommerce as “the procurement of indirect goods or non-production or non-essential goods, including MRO (maintenance, repair, and operations), capital expenditures, and services, including such specifics as:

- Office equipment and supplies
- Computer and other Information Technology equipment
- Piping, electrical, and HVAC supplies
- Legal services
- Utilities and communications
- Catering and food services
- Material handling supplies
- Nuts, bolts, nails, and screws
- Advertising and marketing supplies/services

Historically, these goods have been procured manually via phone, fax, and traditional mail.

• B2C

Business to Consumer is seen mainly as retailing or consumer services aimed to build stronger relationships between the provider of goods and services.

• EProcurement

eProcurement uses Internet and Web technologies to simplify the purchase of these supplies, which are separate from production supplies–the procurement, receipt, and payment of goods for manufacturing or production. According to a Purchasing Magazine estimate, corporate America spends $1.4 trillion on non-production goods annually, which amounts to roughly 20 percent of the US gross domestic product. An eProcurement solution provides a Web-based interface to electronic catalogues inside or outside the firewall, allowing employees to order the goods and services necessary to do their jobs. eProcurement attempts to automate as much of the requisition and procurement process as is possible. This includes the use of the Internet and/or EDI for transactions historically processed manually. The source of the goods or services may be:

• Direct from a manufacturer or service provider
• Through a distributor
• Through a trading network serving as an intermediary

3.2 Subsets

Applications can be divided into three categories:

1. Internal business systems:

• customer relationship management
• enterprise resource planning
• document management systems
• human resources management

2. Enterprise communication and collaboration:

• VoIP
• content management system
• e-mail
• voice mail
• Web conferencing
• Digital work flows (or business process management)

3. Electronic commerce - business-to-business electronic commerce (B2B) or business-to-consumer electronic commerce (B2C):

- internet shop
- supply chain management
- online marketing

3.3 Models

When organizations go online, they have to decide which e-business models best suit their goals. A business model is defined as the organization of product, service and information flows, and the source of revenues and benefits for suppliers and customers. The concept of e-business model is the same but used in the online presence. The following is a list of the currently most adopted e-business models:

- E-shops
- E-procurement
- E-malls
- E-auctions
- Virtual Communities
- Collaboration Platforms
- Third-party Marketplaces
- Value-chain Integrators
- Value-chain Service Providers
- Information Brokerage

3.4 What the Consultants say today

The continuing decline in cost of access to the Internet, rapid improvements in Internet infrastructure and improved secure transaction capability has seen a dramatic rise in E-commerce. The Forrester Group predicts that business-to-business Internet e-commerce will grow rapidly, from $8 billion to $327 billion in goods and services by 2002. On the consumer side the numbers are equally impressive: studies report that the number of Internet users in the United States and Canada doubled in the 18 months prior to March 1997, and that 23 percent of the population over age 16 had used the Internet within the last month. The user population could reach 200 million by 2000.

These changes will have a dramatic effect on both business-to-consumer and business-to-business trading patterns. The dollar volume of Internet-based electronic commerce will expand quickly from an estimated $3 billion in 1996 to $100 billion in 2000, according to an IDC forecast.

New pricing and business models are emerging while the range of available products and services expands every day.

e-Business early adopters are already seeing substantial benefits from e-Business. Although many companies are currently adopting e-Business strategies as a short-term attempt to reduce costs, the most powerful long-term benefits will come through improved relationships between customers and suppliers

So how do you become an e-business company? How can becoming an e-business help you maximize the value of your information technology investment? How can it help you reduce your costs and grow your revenue? There are four important areas or stages in this process. We think of these four stages collectively as the e-business cycle. They are:

- Transforming core business processes.
- Building flexible, expandable e-business applications.
- Running a scalable, available, safe environment.
- Leveraging knowledge and information you've gained through e-business systems.

There is not a set order or hierarchy to this cycle. Successful e-businesses start at different points, and you can, too. But first, you must identify which of your core business processes are most suitable for, or most in need of, conversion to e-business.

Disintermediation

Restricting e-business solutions for commerce to address only transactions has forced businesses to treat Internet e-commerce as a direct-marketing channel strategy, creating conflicts with their other channels and selling partners, damaged relationships, and negative ROI.
E-commerce vendors unable to offer truly interactive solutions have made a virtue out of this necessity. They have claimed that this disintermediation --- the elimination of intermediaries (distributors and resellers, for example) in the chain --- is an advantage of e-commerce.
Disintermediation, they claim, will improve profits and open new markets for businesses that can sell their products and services on the Internet. And they cite numerous examples: Airlines will bypass travel agents and sell direct. Auto makers will eliminate their dealer networks by letting buyers configure their dream cars online.
Based on the disintermediation hype, it would be easy to conclude that in order to succeed at Internet e-commerce, a company must abandon its current business strategy and selling partners. But it just isn't so. Disintermediation is not an opportunity for most businesses, it is a disaster: Adding a transaction server to a company's Web site does not

ensure success. The price of buying and maintaining transaction systems has been high, and the overall cost of producing a public Web site — acquiring and configuring hardware and software, developing custom applications, and building and maintaining content and catalogues — even higher. But most importantly, a transaction server by itself doesn't integrate Internet-based virtual transactions with the way the company does business. And when this new channel isn't integrated into the company's existing systems, the costs of sales can make even successful operations unprofitable. The gold-rush mentality has often meant that if a business didn't put up a Web site, its selling partners would — and the business might find that it had lost control, and its brand had been tarnished in the marketplace. There is no such thing as a free lunch. Intermediaries exist in a selling chain because they add value to the sale. If they are eliminated, their contribution will have to be made up from somewhere else — almost always out of the pocket of the seller.

Common sense -- and the analysts -- argue against disintermediation. Forrester Research says in a February, 1997 study:(6) "Indirect channels won't disappear: Reborn as Internet intermediaries with new capabilities and evolved roles, they will increase in importance as . . . practices shift to acclimate to the new environments." Those words were written about the software industry, but they apply to any industry.

If Internet e-commerce is going to succeed for the enterprise, it must support the business strategy of the enterprise. It cannot be isolated. It cannot be restricted to just handling virtual transactions on a Web site. "Disintermediation" is just a way of narrowing down the vision, of settling for less, of saying that the whole job is too hard, so doing just part of it is good.

Value Chain Integration

Internet e-commerce offers tremendous opportunity for smart intermediaries. These enterprises will figure out new ways of adding value to the selling chain, and they will embrace ways of growing their business by fostering what Forrester Research calls "communities of commerce."

The relationships of the selling chain will continue to be critical to any company's business strategy. To capitalize on them and create communities of commerce, businesses must find Internet e-commerce solutions that support the roles their selling partners fill in the real-world sales cycle:

1. Attracting customers: prospecting, advertising, and informing customers of new products and services.

2. Transacting business: conducting exchanges in ways the customers find comfortable and complete.
3. Retaining customers: Focusing on customer service and relationship-building activities.
4. Managing the process: improving efficiencies, reducing costs, and sharing the benefits among all the parties.

For customers, the benefits of these solutions might include shopping convenience, immediate delivery, more frequent updates, access to more products, and better pricing. Partners in selling and supply chains would benefit from improved communication, access to new markets, a broader product offering, lower cost of doing business, and new ways of adding value. And for a manufacturing business it might mean a lower cost of sales, access to niche markets, better customer identification, and new business models.

The Internet is a massive network that not only links businesses to consumers, but links millions of businesses worldwide to other businesses. There is a clear need for ways to exploit the Internet as a commercial vehicle that doesn't break existing business models, but enables them.

For an Internet e-commerce solution is to succeed it must cascade the benefits of the Internet through the entire chain, from supplier to manufacturer to distributor to reseller to consumer.

3.5 E-Business and E-Commerce

E-business is a well established management term. From the definition (IBM), it can be seen that e-business includes e-commerce, but is broader in scope in that it also refers to the use of Internet technology to support internal process.

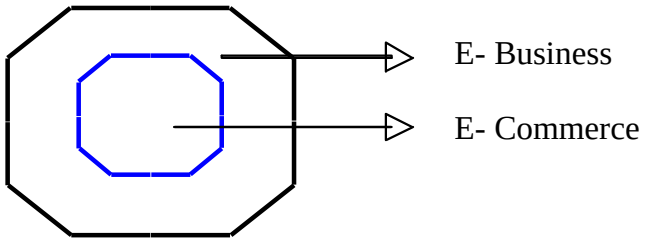


Figure 1: Relationship between E-business and E-commerce

A subset of e-business is e-commerce, which describes the buying and selling of products, services, and information or making transactions via computer networks, including the Internet. The main difference between them is that e-commerce defines interaction between organizations and their customers, clients, or constituents. On the other hand, e-business is

also encompasses an organization's internal operations. In other words, these two can be used interchangeably.

The term E-business was initially crafted in a thematic campaign by IBM in 1996 and subsequently defined as "a secure, flexible, and integrated approach to delivering differentiated business value by combining the systems and processes that run core business operations with the simplicity and reach made possible by Internet technology" (http://www.ibm.com). Prior to the offering of this definition, the term e-business and E- commerce were often referred to interchangeably. The offering of this formal definition marked the coming of age of the adoption of the Internet and its technology to go beyond the function of e-commerce and encompass other functionalities such as e-marketing, e-franchising, e-mailing and many more. In a nutshell, e-business is the function of deploying technology to maximize customer value while e-commerce is the function of creating exchange (i.e., buying and selling) over digital media (Kalakota and Robinson 1999).

3.6 E-Business Concepts

3.6.1 Dimensions of E-Business

Based on the above mentioned perspectives and the degree of digitization of product, process, and the delivery agent, the business can be pure or partial e-business / e-commerce. In traditional commerce, all dimensions are physical while all dimensions are digital in pure e-business / e-commerce. Obviously, in partial e-business / e-commerce, all other possibilities include a mix of digital and physical dimensions. Particularly in the developing countries, the partial e-business / e-commerce have been adopted due to inadequate enabling environment (such as a suitable infrastructure, policies, and financial resources).

3.6.2 Evolution of E-Business/E-Commerce

E-business, that we know of today, has been around for a little over ten years. However, its predecessors, such as Electronic Data Interchange (EDI), Material Requirements Planning (ERP) and Enterprise Resource Planning (ERP), have been around for more than 40 years and still living. EDI is the electronic communication of business transactions, such as orders, confirmations and invoices, between organizations. Although interactive access may be a part of it, EDI implies direct computer to computer transactions into vendors' databases and ordering systems. The detailed information of the MRP and ERP is covered in Module 5. Though, the predecessors cannot rival the exponential growth and acceptance of transacting business over the Internet. In 1970s,

innovations like electronic funds transfer (EFT), funds routed electronically from one organization to another (limited to large corporations), created a new way of doing business or making transactions. In 1990s, the Internet commercialized and users flocked to participate in the form of dot-coms, or Internet start-ups. 1997 saw the introduction of a brand new phrase – e-business, it was seen as a step on from e-commerce (Amazon.com established an e-commerce web site in 1995) and referred to more than just buying and selling via the Internet. In 1999, the emphasis of e-business shifted from B2C to B2B. Then in 2001, the emphasis again shifted from B2B to B2E, c-commerce, e-government, e-learning, and m-commerce. By looking at the trends, e-business / e-commerce will undoubtedly continue to shift and change.

3.7 Rules Governing E-Business

As recognized above, the new paradigm of e-business that is being currently defined is simply technology driven. This changes everything. Kalakota and Robinson map this dramatic paradigm shift by presenting the following as the rules governing e-business:

Rule 1: **Technology in no longer an afterthought in formulating** business strategy, but the actual cause and driver.

Rule 2: **The ability to streamline the structure, influence, and** control of the flow of information is dramatically more powerful and cost-effective than moving and manufacturing physical products

Rule 3: **Inability to overthrow the dominant, outdated business** design often leads to business failure

Rule 4: **The goal of new business designs is to create flexible** outsourcing alliances between companies that not only off-load costs, but also make customers ecstatic

Rule 5: **E-commerce is enabling companies to listen to their** customers and become either “the cheapest,” “the most familiar,” or “the best.”

Rule 6: **Don’t use technology just to create the product. Use** technology to innovate, entertain, and enhance the entire experience surrounding the product, from selection, and ordering to receiving and service.

Rule 7: The business design of the future increasingly uses reconfigurable e-business community models to best meet customer's needs

Rule 8: The tough task for management is to align business strategies, processes, and applications fast, right, and all at once. Strong leadership is imperative.

4.0 CONCLUSION

As a new and emerging concept business, e-business need to be clearly distinguished from other forms of electronic commerce, especially, e-commerce. Proper understanding of the concept goes a long way to determine to what extent it can be utilized as a competitive advantage.

5.0 SUMMARY

This unit is summarized as follows:

- Electronic Business, commonly referred to as "eBusiness" or "e-**Business", may be defined as the utilisation of information and** communication technologies (ICT) in support of all the activities of business.
- There are many definitions for e-business, e-commerce as well as related complementary terms. Some of these definitions have been discussed above. When people hear e-business, they generally think selling over the Web.
- When organizations go online, they have to decide which e-business models best suit their goals. A business model is defined as the organization of product, service and information flows, and the source of revenues and benefits for suppliers and customers.
- The continuing decline in cost of access to the Internet, rapid improvements in Internet infrastructure and improved secure transaction capability has seen a dramatic rise in E-commerce
- E-business is a well established management term. From the definition (IBM), it can be seen that e-business includes e-commerce, but is broader in scope in that it also refers to the use of Internet technology to support internal process.
- Based on the above mentioned perspectives and the degree of digitization of product, process, and the delivery agent, the business can be pure or partial e-business / e-commerce.
- The evolution of e-business / e-commerce became more prominent with the Internet, particularly the World Wide Web or web, revolution.
- As recognized above, the new paradigm of e-business that is being currently defined is simply technology driven.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention 5 perspectives to the definition of e-business.
2. Briefly discuss the rules that guide e-business.

7.0 REFERENCES/FURTHER READINGS

Answers.com, www.answers.com/ e-business, e-business Definitions

Buyitnet.org, www.buyitnet.org/Best_Practice_Guidelines/E-business /index.jsp, E-business

Choi/Stahl/Whinson, (1997). Economics of Electronic Commerce.

Elms, Janelle, (2005). The Seven Essential Steps to Successful eBay *Marketing.*

Gendron, Michael P., (2006). Creating the New e-business Company: *Innovative Strategies for Real-World Applications.*

Porter, Michael E., (2001). Harvard Business Review, Strategy and the *Internet.*

UNIT 2 INTRODUCTION TO E-BUSINESS SECURITY

CONTENTS

1.0 INTRODUCTION

There is a lot of discussion these days about e-business security as more people use email, and services such as banking, mail orders and subscriptions become available through the Internet. But how secure is the Internet and e-commerce and what is computer security?

Today, business is done with many communication technologies, walk-in retail, mail-order phone, mail-order fax etc. The Web and the Internet are just one another communication medium with its own benefits and disadvantages. The cost for a business to have a world wide presence is the lowest in history with the World Wide Web. Budgets of the 1980's would have cost at least $100,000 per month in expenses, to have a business handling international customers 24 hours a day, 7 days a week. Today, that same budget is closer to $5,000 per month and some even much lower. Yet the quality of service that the customer of these businesses is expecting continues to climb.

With these demands, you need a scaleable sales force, accurate, and secure information exchange, automatic delivery of products, and accurate tracking information for package delivery. In this article we will discuss the issues in constructing a web site that can give you all of this and much more. However, there are some pitfalls to be watchful of. The anonymity of the people buying from you can make you feel like you are talking to Mr. X. You rarely have the chance to

speak directly with your customers. It is also far more difficult to get a feel for the size and condition of your vendors and your competitors. You have to help your customers overcome the fear that many people have putting their credit card number into a form on a web page. Using and understanding e-business can give you a strong advantage over your competitors while providing greater value and comfort to your customers.

2.0 OBJECTIVES

At the end of this unit, the student is expected to:

- understand the various types of security measures for e-commerce
- have a basic understanding of how to set up security measures
- specifically know some security measures for a web site
- answer the question of how relevant is security measures
- explain the basic principles for customer security
- understand how to track a customer as a security measure.

3.0 MAIN CONTENT

3.1 Computers and Security

Before the Internet, computer security was limited to ‘closed systems’ or network computers such as offices or banks, where only people physically in the office could use the computer system. It was quite easy for the network supervisor to set up user names and passwords, since at that time people have become used to logging on before they can use these types of computers or resources.

With the advent of the Internet, computers users can now work in an 'open system' and security has become much more complicated. Even though you can now connect your home or office computer to the Internet and perform remote transactions without leaving the building you still want to be sure that the transaction is secure. The transaction takes place through the Internet by bouncing the information through various computers before it reaches, for example, the bank's computer. You want to be sure that no one observes the transaction along the way and collects or modifies your transaction information.

This is where computer security comes in. There are many different types of security systems, though most use a process called encryption. When you connect to your bank or other service to make a transaction you are often required to send your account number or user name as well as a Personal Identification Number (PIN) or password for verification. This information should only be sent after establishing a

secure connection. If you are using an Internet browser you will see a small closed lock appear in the window of the browser. Once you are connected to a secure server any information you send or receive is scrambled or encrypted using a mathematical formula and then reassembled or decrypted at the other end. The computer user usually will not notice this happening as they perform their secure transaction. Anyone with criminal intent who intercepts your transaction will be treated to a stream of garbled nonsense.

If this is the first time you use a new service you most often will need to setup an account and possibly download a small piece of software called a plug-in, which allows your computer to create the secure connection or link.

The transaction often involves the exchange of a small file that keeps track of the transaction and can act as flag or bookmark when you next visit that website. These small files are called cookies and are set by the website you are visiting. They can contain information such as the type of server you are connecting from, the type of browser you are using, the last site you visited and any information you volunteer. You can view the information stored in the cookie. Try a search for 'cookie' to find the cookies folder. Windows users can view any cookies they are storing in the folder

3.2 Security Methods

• Encryption

Privacy is handled by encryption. In PKI (public key infrastructure) a message is encrypted by a public key, and decrypted by a private key. The public key is widely distributed, but only the recipient has the private key. For authentication (proving the identity of the sender, since only the sender has the particular key) the encrypted message is encrypted again, but this time with a private key. Such procedures form the basis of RSA (used by banks and governments) and PGP (Pretty Good Privacy, used to encrypt emails).

Unfortunately, PKI is not an efficient way of sending large amounts of information, and is often used only as a first step — to allow two parties to agree upon a key for symmetric secret key encryption. Here sender and recipient use keys that are generated for the particular message by a third body: a key distribution center. The keys are not identical, but each is shared with the key distribution center, which allows the message to be read. Then the symmetric keys are encrypted in the RSA manner, and rules set under various protocols. Naturally, the private keys have to be kept secret, and most security lapses indeed arise here.

Encryption also involves using the key pair but in reverse. Once your message is completed you encrypt the file using the recipient's public key ensuring that only the recipient can ever access that message with their private key

• Digital Signatures and Certificates

Digital signatures meet the need for authentication and integrity. To vastly simplify matters (as throughout this page), a plain text message is run through a hash function and so given a value: the message digest. This digest, the hash function and the plain text encrypted with the recipient's public key is sent to the recipient. The recipient decodes the message with their private key, and runs the message through the supplied hash function to that the message digest value remains unchanged (message has not been tampered with). Very often, the message is also times tamped by a third party agency, which provides non-repudiation.

What about authentication? How does a customer know that the website receiving sensitive information is not set up by some other party posing as the e-merchant? They check the digital certificate. This is a digital document issued by the CA (certification authority: Verisign, Thawte, etc.) that uniquely identifies the merchant. Digital certificates are sold for emails, e-merchants and web-servers. Digital signature shall be discussed in detail in subsequent units of this course.

• Secure Socket Layers

SSL stands for Secure Sockets Layer. This is the technique in which web servers and web browsers encrypt and decrypt all of the information that they transmit and receive. Secret decoder rings time. Both ends establish and use the same scheme for making sure that no one else is listening to their conversation.

Information sent over the Internet commonly uses the set of rules called TCP/IP (Transmission Control Protocol / Internet Protocol). The information is broken into packets, numbered sequentially, and an error control attached. Individual packets are sent by different routes. TCP/IP reassembles them in order and resubmits any packet showing errors. SSL uses PKI and digital certificates to ensure privacy and authentication. The procedure is something like this: the client sends a message to the server, which replies with a digital certificate. Using PKI, server and client negotiate to create session keys, which are symmetrical secret keys specially created for that particular

transmission. Once the session keys are agreed, communication continues with these session keys and the digital certificates.

• PCI, SET, Firewalls and Kerberos

Credit card details can be safely sent with SSL, but once stored on the server they are vulnerable to outsiders hacking into the server and accompanying network. A PCI (peripheral component interconnect: hardware) card is often added for protection, therefore, or another approach altogether is adopted: SET (Secure Electronic Transaction). Developed by Visa and Mastercard, SET uses PKI for privacy, digital certificates to authenticate the three parties: merchant, customer and bank. More importantly, sensitive information is not seen by the merchant, and is not kept on the merchant's server.

Firewalls (software or hardware) protect a server, a network and an individual PC from attack by viruses and hackers. Equally important is protection from malice or carelessness within the system, and many companies use the Kerberos protocol, which uses symmetric secret key cryptography to restrict access to authorized employees.

3.3 Setting Up Security

As most people will not be setting up their own secure server the scope of this section is limited to the topics of protecting e-mail and small business or organizational transactions.

E-mail can be protected using a service or an application (program). There are others but the two that stand out currently are S/MIME and **PGP. S/MIME requires the user to register with a 3 party service which** issues a digital id that you attach to your message. Though this is usually a commercial service there is often a free introductory period. PGP is free for personal use or a commercial application for business use and is run from your own computer.

Both methods allow users to sign or attach a digital identification to the email message which verifies, to the recipient, that the message is from the original person or organization and that the information wasn't tampered with in transit. These methods also allow the user to encrypt their message so that anyone intercepting the message wouldn't be able to read it. You can also decide the level of encryption from low; in which a nerd with some good software and enough time on their hands could possibly decrypt to high (128 bit) which would take a whole team of experts weeks to decrypt if even then. Most of us will choose somewhere in between as this process involves increased time and file size.

Both methods use key pairs of public and private keys. Your public keys are sent to everyone that you communicate through email with. Your public key can be sent through various methods including posting it to an internet service or sending it as part of an email message. Public keys can also be post on your website in a file. Your friends and associated can add your public key to a file called a key ring). When someone wants to send you a secure email the sender encrypts their messages with your public key. When you receive the email you must decrypt it using your private key. Many email programs will automatically verify that the message is authentic. You will need to type in your password to view the message.

Small businesses and organizations that wish to offer transactions over the Internet or e-business can take their chances and set up an unsecured system, set up their own secure server or purchase a service from a third party. There are various types including service that take a percentage of the transaction and/or charge a service fee and/or charge for each transaction. Some organizations are more reliable and you should always shop around before committing to a service. Because this type of service is so new the length of time a company has been operating is not always a way to decide. Things to watch for are downtime. If your company's website is operating properly yet the customer or user can't access the transaction server because it is down, too busy or misconfigured they will easily be put off perhaps entirely. Watch for contracts that lock you in as the market is still developing and prices tend to fluctuate. It is easy to switch services by simply changing the address on your website's order forms.

3.4 Security and Websites

As stated at the beginning of this unit, the nature of the Internet is an open system. Having said that, there are many reasons and many ways to set up a secure or closed system within this open framework. Private or member-based discussion groups, private files or folders, protected databases, copyright material to name a few all need some way of allowing them to be distributed to the intended recipient only. Also, many businesses are creating Intranets which are closed systems only accessible to registered users. An Intranet can provide a way of making company information easily accessible and allow branch offices to communicate with each other easier.

Account Security

Your website itself is protected by your ISP's software. When you attempt to access your web space to change or modify a file using a shell or ftp you are challenged to send your username and password. This is the first line of protection and adequate for many website administrators.

Server Security

The server that your website is installed on is the second line of protection. Most servers have security features built in to them allowing users to password protect folders or build scripts to send a username/password challenge to a user trying to access a file or folder. This allows website administrators the ability to create discussion groups within their site or to place confidential documents or information that is made available only to registered users on their own website. Unfortunately some ISP either don't make this option available, charge a premium to use them or only allow their own employees to set them up.

Third Party Security

Another option includes contracting the protection of private files to a separate service, pay a third party to hosting a private discussion group or obtain web space on another server that allows access to security options. The entire Internet is as close as your computer connection and whether the file the user is viewing is stored in your own current web space or on another server is usually immaterial. When your customers, employees or members moves from one page to another the consistency of the website is the maintained by the design, not the address of the separate pages. It is also possible to control the address that is displayed if required.

Software Security

Another option is to use JavaScript or Java applets to control how customers or members access secure features. This option is only available to users who are using Java enabled browsers. Scripts and applets can control access to documents and databases, create content on the fly based on user input, detect the browser the visitor is using and direct them to the proper page, retrieve cookies and use that information to determine whether a user has access to a certain area or not, as well as many other uses.

Copyright

Copyright is a protect using the same process as any original material (books, artwork, film, etc...). Anything that a user gets off the Internet should be treated as privately owned information unless otherwise noted. Anyone posting private information to the Internet should be aware that copyright law is not the same in every country and may be difficult to enforce. It is possible to set up a page that won't be stored on the users computer once they leave the site but that will only slow down not stop users who want to obtain information posted on a website. Notices of copyright are often added to the main page of a website sometimes with a link to a page describing the details of how the content can be used.

Updating Software

It is very important to update your software periodically. When a program is released, particular internet browsers, it may contain flaws usually referred to as bugs. These bugs may not appear to be a problem but criminals will attempt to use these flaws for their own use. Keeping your software up to date will help keep your computer secure.

3.5 Is Security Necessary?

Though you may think that it is not necessary to setup security systems there are many reason to consider it. I have come across a number of examples of people forging documents and email. A digital signature will be the only way to verify whether a document is genuine or not.

Many organizations need to discuss draft articles, changes to bylaws and other documents that could cause problems if they were made public before they are approved. A secure directory within your website is an ideal spot to store sensitive material making it available for members and people who have the proper password.

I would be remiss to not point out and as all discussions on the subject also point out mining the Internet with malicious intent is also possible. One common malicious act is to search websites for email addresses and then add them to spam distribution lists. Unfortunately there is very little that can be done to counter this other than removing your email address from your web site but this makes it difficult for your customers to contact you.

Whether you decide to add a security component to your web site project initially it is a good idea to think about or have a discussion about web site security when planning the site. You should also review your security systems periodically whether that is changing your password or reviewing and updating your security system.

3.6 Customer Security: Basic Principles

Most e-commerce merchants leave the mechanics to their hosting company or IT staff, but it helps to understand the basic principles. Any system has to meet four requirements:

- Privacy: information must be kept from unauthorized parties.
- Integrity: message must not be altered or tampered with.
- Authentication: sender and recipient must prove their identities to each other.
- Non-repudiation: proof is needed that the message was indeed received.

Transactions

Sensitive information has to be protected through at least three transactions:

- credit card details supplied by the customer, either to the merchant or payment gateway. Handled by the server's SSL and the merchant/server's digital certificates.
- credit card details passed to the bank for processing. Handled by the complex security measures of the payment gateway.
- order and customer details supplied to the merchant, either directly or from the payment gateway/credit card processing company. Handled by SSL, server security, digital certificates (and payment gateway sometimes).

3.7 Practical Consequences

1. The merchant is always responsible for security of the Internet-connected PC where customer details are handled. Virus protection and a firewall are the minimum requirement. To be absolutely safe, store sensitive information and customer details on zip-disks, a physically separate PC or with a commercial file storage service. Always keep multiple back-ups of essential information, and ensure they are stored safely off-site.

2. Where customers order by email, information should be encrypted with PGP or similar software. Or payment should be made by specially encrypted checks and ordering software.

3. Where credit cards are taken online and processed later, it's the merchant’s responsibility to check the security of the hosting company’s webserver. Use a reputable company and demand

detailed replies to your queries.

4. Where credit cards are taken online and processed in real time, four situations arise:

i. You use a service bureau. Sensitive information is handled entirely by the service bureau, which is responsible for its security. Other customer and order details are your responsibility as in 3 above.

ii. You possess an e-business merchant account but use the digital certificate supplied by the hosting company. A cheap option acceptable for smallish transactions with SMEs. Check out the hosting company, and the terms and conditions applying to the digital certificate.

iii. You possess an e-business merchant account and obtain your own digital certificate (costing some hundreds of dollars). Check out the hosting company, and enter into a dialogue with the certification authority: they will certainly probe your credentials.

iv. You possess a merchant account, and run the business from your own server. You need trained IT staff to maintain all aspects of security-firewalls, Kerberos, SSL, and a digital certificate for the server (costing thousands or tens of thousands of dollars).

Security is a vexing, costly and complicated business, but a single lapse can be expensive in lost funds, records and reputation. Do not wait for disaster to strike, but stay proactive, employing a security expert where necessary.

Sites on our resources page supplies details.

3.8 Tracking the Customer

Of primary importance in any transaction is that the customer feels comfortable with your communication. To make it seem like the website is talking to each customer individually you must track who the customer is and what he is interested in. The most common way this is achieved on the web is with the shopping cart concept. This allows many different people to be shopping on your site and all have their own sets of items in their cart. In our fax back example you would have to use something like the fax number to keep track of each customer. The equivalent with the web would be the IP number (known as IP tracking). The one major difference is that a customer's fax number doesn't change very often, while a customer's IP number can change every time that

they connect to the Internet -- for those people using dial up accounts or other dynamic addressing situations -- so IP numbers are not a very reliable way to track customers.

Another common tracking technique is cookies. You can have your website put a cookie onto the customer's machine so that it maintains important information, like the contents of their shopping cart. A better technique that I have found is tag propagation. This is a technique in which the first page that someone hits when they enter the site assigns a unique number, something like the number of seconds since 1904. This number is in turn passed thru every page on the site and the shopping cart information is stored in a file with that number on the server. This allows a customer to disconnect (by choice or happenstance) from the Internet and not loose the shopping cart information. This can be very important in situations where buying approval from someone else is required for the purchase. Most of the commercial products include a way of doing this. With WebCatalog you insert a cart=[cart] parameter into every HREF and form on your site.

Tracking the customer is very useful not just for the convenience of a shopping cart, but for things like tracking down people that you think are using stolen cards and, more importantly for that all allusive goal, to make the site more usable. Correlating this tracking information with the general web server logs can be used to determine trends of the people visiting your site, are they getting all the information they need to make a buying decision, are they understanding the buying process, are they loosing interest after a certain amount of time. One big advantage of this tracking log is to look for all the searches that people are doing on your site and were they are not finding any products. Maybe you should be the products more effectively. All of these answers can help you understand ways to change your site to make it more useful.

3.9 Security Concerns

Areas That Need Security

As mentioned in the section about SSL, we do want to protect the transmission of sensitive information with something like SSL to keep the eavesdroppers away, but another equally important issue for security is protection from attacks on your web server. People trying to find credit card numbers in accounting logs or just trying to steal products, to buy at ridiculously low or free prices. Prevention of this type of security breach is the most overlooked area. Much of the information on the machine should not be allowed any access. You don't want people knowing even about access statistics without you knowing about it.

The first obvious area to secure is the accounting files. Let's say the web server is doing a great job of keeping people out of sensitive areas, but the same machine is also your ftp server. People are prevented by the web server from getting to your accounting log, but maybe there is a security hole because your ftp server software allows access to this log... so my first advice, limit the access protocols to all sensitive data -- 1) store your accounting logs and other sensitive files outside of the web server folder, WebStar and many other web server products will not serve files outside of their folder tree, 2) don't run ftp and other protocol services on the same machine. Also, make sure that if you are delivering electronic product, only the person that bought it, gets it. For this you should either be copying the product to some unique place only that person is given access to or have a one time password scheme allowing only one shot at downloading the product.

The concern of the web server allowing access to files that are sensitive is best taken care of by your disk organization. Below is a screen shot of a sample organization of your web server folder structure using WebStar and WebCatalog:

Areas That Do Not Need Security

There are many areas within the selection and buying process that are considered public information and therefore don't need security. In fact, the whole process would be slowed down if it sent everything through a SSL server. Imagine if you received a mail-order catalog from MacWarehouse or Club-Mac and you had to put a decoder ring over each letter to figure out what it really was, that would take you hours just to read one page. That is what your browser is doing with SSL data. So, big picture, you only want to use SSL when you are expecting sensitive data from the customer, like a credit card number. Protect that from eavesdroppers with SSL, everything else should go thru the non-SSL server.

4.0 CONCLUSION

E-Commerce is more secure than most business we conduct everyday and is getting better every minute. Knowing various hacking techniques on the Internet and having built an e-Commerce package, if I wanted to get a few credit card numbers I would head for the local bar and go thru the dumpster long before I would start going after websites. Give yourself time to understand and work with your new sales force. A properly constructed website benefits the consumer with up to the minute information and immediate response. The same website serves as hundreds of sales people for the merchant, all trained with exactly the right information as well as access to tracking information etc. The

positive return for the customer and the merchant will help to overcome the myth and fear of the security on the Internet. I would like to end on a observation about most credit cards, even if it is stolen, the owner is only liable for $50.

- There are a variety of tools on the market to help you construct your eCommerce web site. Each has its own strengths and weaknesses. To choose the best for your needs, you must carefully research the speed and responsiveness of the server under load, how they handle the security areas and your database connectivity needs, do they have to handle a live existing database.

5.0 SUMMARY

- There is a lot of discussions these days about e-commerce security as more people use email and more services such as banking, orders and subscriptions become available through the Internet
- Digital signatures meet the need for authentication and integrity. To vastly simplify matters (as throughout this page), a plain text message is run through a hash function and so given a value: the message digest
- Email can be protected using a service or an application (program). There are others but the two that stand out currently are S/MIME and PGP. S/MIME requires the user to register with a 3 party service which issues a digital id that you attach to your message
- As was stated at the beginning of this unit, the nature of the Internet is an open system. Having said that there are many reasons and many ways to set up a secure or closed system with in this open framework
- Though you may think that it is not necessary to setup security systems, there are many reasons to consider it. I have come across a number of examples of people forging documents and email. A digital signature will be the only way to verify whether a document is genuine or not.
- Before the Internet, computer security was limited to 'closed systems' or network computers such as offices or banks where only people physically in the office could use the computer system. It was quite easy for the network supervisor to set up user names and passwords and since that time people have become used to logging on before they can use these types of computers or resources.
- Most e-commerce merchants leave the mechanics to their hosting company or IT staff, but it helps to understand the basic principles. Any system has to meet four requirements
- Of primary importance in any transaction is that the customers feel comfortable with your communication. To make it seem like the website is talking to each customer individually you must track who the customer is and what they are interested in

• The first obvious area to secure is the accounting files. Let's say the web server is doing a great job of keeping people out of sensitive areas, but the same machine is also your ftp server

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention 5 security methods for e-commerce security
2. Mention 4 basic requirements of any system to secure a customer within the context of this unit.

7.0 REFERENCES/FURTHER READINGS

Jay, Van Vark. eCommerce and the Security Myth.

UNIT 3 E-BUSINESS SECURITY CHALLENGES

CONTENTS

1.0 INTRODUCTION

The new millennium brought with it new possibilities in terms of information access and availability, simultaneously introducing new challenges in protecting sensitive information from some eyes while making it available to others. The Internet allows businesses to use information more effectively, by allowing customers, suppliers, employees, and partners to get access to the business information they need, when they need it. These Internet-enabled services all translate to reduced cost: there is less overhead, greater economies of scale, and increased efficiency. E-business' greatest promise is more timely, more valuable information accessible to more people, at reduced cost of information access. With the changes in business operations as a result of the Internet era, security concerns move from computer labs to the front page of newspapers. The promise of e-business is offset by the security challenges associated with the disintermediation of data access. One security challenge results from "cutting out the middleman," that too often cuts out the information security the middleman provides. Another is the expansion of the user community from a small group of known, vetted users accessing data from the intranet, to thousands of users accessing data from the Internet. Application service providers (ASP) and exchanges offer especially stringent—and sometimes contradictory—requirements of per user and per customer security, while allowing secure data sharing among communities of interest.

E-business depends on providing customers, partners, and employees with access to information, in a way that is controlled and secure. Technology must provide security to meet the challenges encountered by e-businesses. Virtually all software and hardware vendors claim to build secure products, but what assurance does an e-business have of a product's security? E-businesses want a clear answer to the conflicting security claims they hear from vendors. How can you be confident about the security built into a product? Independent security evaluations against internationally-established security criteria provide assurance of vendors' security claims.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

• explain the greatest promises of e-business

• understand the needs for the security of e-business
• discuss cases of problems encountered in e-business transactions.

3.0 MAIN CONTENT

3.1 The Needs of E-Business Security

While putting business systems on the Internet offers potentially unlimited opportunities for increasing efficiency and reducing cost, it also offers potentially unlimited risk. The Internet provides much greater access to data, and to more valuable data, not only to legitimate users, but also to hackers, disgruntled employees, criminals, and corporate spies.

Increased Data Access
One of the chief e-business benefits of the Internet is "disintermediation." The intermediate information processing steps that employees typically perform in "brick and mortar" businesses, such as typing in an order received over the phone or by mail, are removed from the e-business process. Users who are not employees and are thus outside the traditional corporate boundary, including customers, suppliers, and partners, can have direct and immediate online access to business information which pertains to them.

In a traditional office environment, any access to sensitive business information is through employees. Although employees are not always reliable, at least they are known, their access to sensitive data is limited by their job function, and access is enforced by physical and procedural controls. Employees who pass sensitive information outside the company contrary to policy may be subject to disciplinary action; the threat of punishment thus helps prevent unauthorized access. Making business information accessible via the Internet vastly increases the number of users who may be able to access that information. When business is moved to the Internet, the environment is drastically changed. Companies may know little or nothing about the users (including, in many cases, employees) who are accessing their systems. Even if they know who their users are, it may be very difficult for companies to deter users from accessing information contrary to company policy. It is therefore important that companies manage access to sensitive information, and prevent unauthorized access to that information before it occurs.

Much More Valuable Data

E-business relies not only on making business information accessible outside the traditional company, it also depends on making the best,

most up-to-date information available to users when they need it. For example, companies can streamline their operations and reduce overhead by allowing suppliers to have direct access to consolidated order information. This allows companies to reduce inventory by obtaining exactly what they need from suppliers when they need it. Streamlining information flow through the business system allows users to obtain better information from the system. Now, businesses that allow other businesses and consumers to submit and receive information directly through the Internet can expect to get more timely, accurate, and valuable information, at less expense than if traditional data channels were used.

Formerly, when information was entered into a business system, it was often compartmentalized. Information maintained by each internal department, such as sales, manufacturing, distribution, and finance, was kept separate, and was often processed by physically separate and incompatible databases and applications—so-called "islands of information." Companies have found that linking islands of information and consolidating them where possible, allows users to obtain better information, and to get more benefit from that information, which thus makes the information more valuable.

Improving the value of data available to legitimate users generally improves its value to intruders as well, increasing the potential rewards to be gained from unauthorized access to that data, and the potential damage that can be done to the business if the data were corrupted. In other words, the more effective an e- business system is, the greater the need to protect it against unauthorized access.

Scalability with Large User Communities

The sheer size of the user communities which can access business systems via the Internet not only increases the risk to those systems, it also constrains the solutions which can be deployed to address that risk. The Internet creates challenges in terms of scalability of security mechanisms, management of those mechanisms, and the need to make them standard and interoperable. Security mechanisms for Internet-enabled systems must support much larger communities of users than systems that are not Internet-enabled. Whereas the largest traditional enterprise systems typically supported thousands of users, many Internet-enabled systems have millions of users.

Manageability

Traditional mechanisms for identifying users and managing their access, such as granting each user an account and password on each system he

accesses, may not be practical in an Internet environment. It rapidly becomes too difficult and expensive for system administrators to manage separate accounts for each user on every system.

Interoperability

Unlike traditional enterprise systems, where a company owns and controls all components of the system, Internet-enabled e-business systems must exchange data with systems owned and controlled by others: customers, suppliers, partners, etc. security mechanisms deployed in e-business systems must therefore be standards based, flexible, and interoperable, to ensure that they work with others' systems. They must support browsers, and work in multi-tier architectures with one or more middle tiers such as web servers and application servers.

Hosted Systems and Exchanges

The principal security challenge of hosting is keeping data from different hosted user communities separate. The simplest way of doing this is to create physically separate systems for each hosted community. The disadvantage of this approach is that it requires a separate computer, with separately installed, managed, and configured software, for each hosted user community, providing little economies of scale to a hosting company. Mechanisms that allow multiple different user communities to share a single hardware and software instance, keep data for different user communities separate, and allow a single administrative interface for the hosting provider, can greatly reduce costs for the hosting service provider. Exchanges have requirements for both data separation and data sharing. For example, an exchange may ensure that a supplier's bid remains unviewable by other suppliers, yet allow all bids to be evaluated by the entity requesting the bid.

Furthermore, exchanges may also support "communities of interest" in which groups of organizations can share data selectively, or work together to provide a joint bid, for example. Assurance

E-businesses need some form of assurance of the security provided in the technology products they purchase. For such assurance, there are international standards used to validate vendors' security claims against established criteria in formal evaluations.

Security evaluations are carried out by independent, licensed and accredited organizations. The evaluation process, from inception to certificate, often lasts up to a full year (and sometimes longer). Vendors who have undergone evaluations of their products learn to improve upon

their development, testing and shipping processes as a result of completing the demanding process. Security evaluations are perhaps the most effective way to qualify a vendor's assertions about its security implementations. Is a product that has not completed such evaluations secure enough to run an e-business? Is it secure enough to protect an organization's most sensitive data? E-businesses demand that the software and hardware vendors they select ship certified provably secure products. Assurance afforded by independent security evaluations lets e-businesses be assured of the products they purchase and deploy.

3.2 Information Systems Breakdowns

Businesses that depend on computer face lots of threats and breakdown. The following incidents and cases illustrate representative cases of breakdowns in the information systems of e-businesses.

Incident 1

For almost two weeks in 1993, a seemingly legitimate automated teller machine (ATM) operating in a shopping mall near Hartford, Connecticut gave consumers apologetic notes that said "sorry, no transactions are possible". Meanwhile the machine recorded the card numbers and the personal identification numbers that hundreds of customers entered in their vain attempt to make the machine dispense cash. On May 8, 1993, while the dysfunctional machine was still sitting in the shopping mall, thieves started tapping into the 24-hour automated teller network in New York City. Using counterfeit bank cards encoded with the numbers stolen from the Hartford customers, the thieves removed almost $100,000 dollars from the accounts of innocent customers. The criminals were successful in making an ATM machine do what it was supposedly not designed to do, breach its own security by recording bank card numbers together with personal security codes.

Incident 2

In 1994, a Russian hacker who did not know English broke into a Citibank electronic fund transfer system and stole more than $10 million by wiring it into accounts around the world. Since then, Citibank, a giant bank that moves half a trillion dollars a day, increased its security measures, requiring customers to use electronic devices that create new passwords very frequently.

Incident 3

According to Wall Street Journal, the Bank of Tokyo–Mitsubishi branch in New York and the National Westminster Bank in the UK reported

losses of tens of millions of dollars in 1996v due to errors in their options and derivatives trading models. I both cases, the losses went undetected for a long time. In the first case the trading model was found to be inaccurate, in the second case the model was fed inaccurate data.

Incident 4

Netscape security is aimed at scrambling sensitive financial data such as credit card numbers and sales transactions so they would be safe from break-ins, by using a powerful 128-bit program. However, using 120 powerful workstations and two supercomputers, in 1996 a French student breached the encryption program in eight days, demonstrating that no program is 100 percent secure.

Incident 5

In 1996, the Los Angeles Times reported "Computer makes $850 million error in Social Security" .The glitch shortchanged about 700,000 Americans in retirement benefits and had been undetected for almost 23 years until it was discovered during an audit in 1994. While the newspaper blamed the computer, the fault is actually that of the programmers who were unable to properly automate the complex computations of the benefits. It took more than three years to fix the problem.

Incident 6

A Tarrant County, TX jury found Donald Burleson guilty of harmful access to a computer, a third degree felony with a maximum penalty of 10 years in prison and a $5,000 fine. Jurors were told that the man planted a virus in a computer system that was used to store records by an insurance and brokerage firm. The virus was programmed like a time bomb and was activated two days after the man was fired from his job. The virus eliminated 168,000 payroll records, which resulted in a one-month delay in issuing employees payroll checks.

Incident 7

In 1999, a fire disabled a major Illinois Bell switching center. The outage affected the voice and data communications of more than one-half million residents and hundreds of businesses during a period ranging from two days to three weeks. The major effects on business were the following:

- Dozens of banks were hindered in cashing checks and transferring funds.
- At least 150 travel agencies were hindered in their ability to make reservations and print tickets
- About 300 automated teller machines were shut down
- Most of the cellular phones and paging systems in the area were disrupted
- Hundreds of companies were hindered in their communications, both inside and outside the immediate area.

4.0 CONCLUSION

E-business depends on providing customers, partners, and employees with access to information, in a way that is controlled and secure. Managing e-business security is a multifaceted challenge and requires the coordination of business policy and practice with appropriate technology. In addition to deploying standards bases, flexible and interoperable systems, the technology must provide assurance of the security provided in the products. As technology matures and secure e-business systems are deployed, companies will be better positioned to manage the risks associated with disintermediation of data access. Through this process businesses will enhance their competitive edge while also working to protect critical business infrastructures from malefactors like hackers, disgruntled employees, criminals and corporate spies.

5.0 SUMMARY

- The new millennium brought with it new possibilities in terms of information access and availability, simultaneously introducing new challenges in protecting sensitive information from some eyes while making it available to others.
- While putting business systems on the Internet offers potentially unlimited opportunities for increasing efficiency and reducing cost, it also offers potentially unlimited risk.
- The principal security challenge of hosting is keeping data from different hosted user communities separate
- Businesses that depend on computer face lots of threats and breakdown.
- For almost two weeks in 1993, a seemingly legitimate automated teller machine (ATM) operating in a shopping mall near Hartford, Connecticut gave consumers apologetic notes that said “sorry, no transactions are possible”.
- In 1999, a fire disabled a major Illinois Bell switching center. The outage affected the voice and data communications of more than

one-half million residents and hundreds of businesses during a period ranging from two days to three weeks.

• E-business depends on providing customers, partners, and employees with access to information, in a way that is controlled and secure.

6.0 TUTOR-MARKED ASSIGNMENT

Briefly discuss the valuing of data as a need for e-business security.

7.0 REFERENCES/FURTHER READINGS

Lord, P. Davidson, M.A. and Browder, K. (2002). Managing e-business *Security Challenges. U.S.A: Oracle Corporation World* Headquarters.

Turban, E. McLean, E. and Wetherbe, J. (1999). Information *Technology for Management. John Wiley & Sons.*

UNIT 4 TYPES OF INFORMATION SECURITY CONTROLS

CONTENTS

1.0 INTRODUCTION

Security is generally defined as the freedom from danger or as condition of safety. Computer security, specifically, is the protection of data in a system against unauthorized disclosure, modification, or destruction and protection of the computer system itself against unauthorized use, modification, or denial of service. Because certain computer security controls inhibit productivity, security is typically a compromise toward which security practitioners, system users, and system operations and administrative personnel work to achieve a satisfactory balance between security and productivity.

Controls for providing information security can be physical, technical, or administrative. These three categories of controls can be further classified as either preventive or detective. Preventive controls attempt to avoid the occurrence of unwanted events, whereas detective controls attempt to identify unwanted events after they have occurred. Preventive controls inhibit the free use of computing resources and therefore can be applied only to the degree that the users are willing to accept. Effective security awareness programs can help increase users' level of tolerance for preventive controls by helping them understand how such controls enable them to trust their computing systems. Common detective controls include audit trails, intrusion detection methods, and checksums.

Three other types of controls supplement preventive and detective controls. They are usually described as deterrent, corrective, and recovery. Deterrent controls are intended to discourage individuals from intentionally violating information security policies or procedures. These usually take the form of constraints that make it difficult or undesirable to perform unauthorized activities or threats of consequences that influence a potential intruder to not violate security (e.g., threats ranging from embarrassment to severe punishment).

Corrective controls either remedy the circumstances that allowed the unauthorized activity or return conditions to what they were before the violation. Execution of corrective controls could result in changes to existing physical, technical, and administrative controls. Recovery

controls restore lost computing resources or capabilities and help the organization recover monetary losses caused by a security violation.

Deterrent, corrective, and recovery controls are considered to be special cases within the major categories of physical, technical, and administrative controls; they do not clearly belong in either preventive or detective categories. For example, it could be argued that deterrence is a form of prevention because it can cause an intruder to turn away; however, deterrence also involves detecting violations, which may be what the intruder fears most. Corrective controls, on the other hand, are not preventive or detective, but they are clearly linked with technical controls when antiviral software eradicates a virus or with administrative controls when backup procedures enable restoring a damaged database. Finally, recovery controls are neither preventive nor detective but are included in administrative controls as disaster recovery or contingency plans.

Because of these overlaps with physical, technical, and administrative controls, the deterrent, corrective, and recovery controls are not discussed further in this chapter. Instead, the preventive and detective controls within the three major categories are examined.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define security
- identify and differentiate the types of controls
- explain the strategies for physical control
- understand the strategies for technical control
- define the strategies for administrative control.

3.0 MAIN CONTENT

3.1 Physical Controls

Physical security is the use of locks, security guards, badges, alarms, and similar measures to control access to computers, related equipment (including utilities), and the processing facility itself. In addition, measures are required for protecting computers, related equipment, and their contents from espionage, theft, and destruction or damage by accident, fire, or natural disaster (e.g., floods and earthquakes).

Preventive Physical Controls

Preventive physical controls are employed to prevent unauthorized personnel from entering computing facilities (i.e., locations housing computing resources, supporting utilities, computer hard copy, and input data media) and to help protect against natural disasters. Examples of these controls include:

- Backup files and documentation.
- Fences.
- Security guards.
- Badge systems.
- Double door systems.
- Locks and keys.
- Backup power.
- Biometric access controls.
- Site selection.
- Fire extinguishers.

Backup Files and Documentation: Should an accident or intruder destroy active data files or documentation, it is essential that backup copies be readily available. Backup files should be stored far enough away from the active data or documentation to avoid destruction by the same incident that destroyed the original. Backup material should be stored in a secure location constructed of noncombustible material, including two-hour-rated fire walls. Backups of sensitive information should have the same level of protection as the active files on the system; it is senseless to provide tight security for data on line but lax security for the same data in a backup location.

Fences: Although fences around the perimeter of the building do not provide much protection against a determined intruder, they do establish a formal no trespassing line and can dissuade the simply curious person. Fences should have alarms or should be under continuous surveillance by guards, dogs, or TV monitors.

Security Guards: Security guards are often stationed at the entrances of facilities to intercept intruders and ensure that only authorized persons are allowed to enter. Guards are effective in inspecting packages or other hand-carried items to ensure that only authorized, properly described articles are taken into or out of the facility. The effectiveness of stationary guards can be greatly enhanced if the building is wired with appropriate electronic detectors with alarms or other warning indicators terminating at the guard station. In addition, guards are often used to patrol unattended spaces inside buildings after normal working hours to deter intruders from obtaining or profiting from unauthorized access.

Badge Systems: Physical access to computing areas can be effectively controlled using a badge system. With this method of control, employees and visitors must wear appropriate badges whenever they are in access-controlled areas. Badge-reading systems programmed to allow entrance only to authorized persons can then easily identify intruders.

Double Door Systems: Double door systems can be used at entrances to restricted areas (e.g., computing facilities) to force people to identify themselves to the guard before they can be released into the secured area. Double doors are an excellent way to prevent intruders from following closely behind authorized persons and slipping into restricted areas.

Locks and Keys: Locks and keys are commonly used for controlling access to restricted areas. Because it is difficult to control copying of keys, many installations use cipher locks (i.e., combination locks containing buttons that open the lock when pushed in the proper sequence). With cipher locks, care must be taken to conceal which buttons are being pushed to avoid a compromise of the combination.

Backup Power: Backup power is necessary to ensure that computer services are in a constant state of readiness and to help avoid damage to equipment if normal power is lost. For short periods of power loss, backup power is usually provided by batteries. In areas susceptible to outages of more than 15–30 min., diesel generators are usually recommended.

Biometric Access Controls: Biometric identification is a more sophisticated method of controlling access to computing facilities than badge readers, but the two methods operate in much the same way. Biometrics used for identification include fingerprints, handprints, voice patterns, signature samples, and retinal scans. Because biometrics cannot be lost, stolen, or shared, they provide a higher level of security than badges. Biometric identification is recommended for high-security, low-traffic entrance control.

Site Selection: The site for the building that houses the computing facilities should be carefully chosen to avoid obvious risks. For example, wooded areas can pose a fire hazard, areas on or adjacent to an earthquake fault can be dangerous and sites located in a flood plain are susceptible to water damage. In addition, locations under an aircraft approach or departure route are risky, and locations adjacent to railroad tracks can be susceptible to vibrations that can precipitate equipment problems.

Fire Extinguishers: The control of fire is important to prevent an emergency from turning into a disaster that seriously interrupts data processing. Computing facilities should be located far from potential fire sources (e.g., kitchens or cafeterias) and should be constructed of noncombustible materials. Furnishings should also be noncombustible. It is important that appropriate types of fire extinguishers be conveniently located for easy access. Employees must be trained in the proper use of fire extinguishers and in the procedures to follow should a fire break out.

Automatic sprinklers are essential in computer rooms and surrounding spaces and when expensive equipment is located on raised floors. Sprinklers are usually specified by insurance companies for the protection of any computer room that contains combustible materials. However, the risk of water damage to computing equipment is often greater than the risk of fire damage. Therefore, carbon dioxide extinguishing systems were developed; these systems flood an area threatened by fire with carbon dioxide, which suppresses fire by removing oxygen from the air. Although carbon dioxide does not cause water damage, it is potentially lethal to people in the area and is now used only in unattended areas.

Detective Physical Controls
Detective physical controls warn protective services personnel that physical security measures are being violated. Examples of these controls include:

- Motion detectors.
- Smoke and fire detectors.
- Closed-circuit television monitors.
- Sensors and alarms.

Motion Detectors: In computing facilities that usually do not have people in them, motion detectors are useful for calling attention to potential intrusions. Motion detectors must be constantly monitored by guards.

Fire and Smoke Detectors: Fire and smoke detectors should be strategically located to provide early warning of a fire. All fire detection equipment should be tested periodically to ensure that it is in working condition.

Closed-Circuit Television Monitors: Closed-circuit televisions can be used to monitor the activities in computing areas where users or operators are frequently absent. This method helps detect individuals behaving suspiciously.

Sensors and Alarms: Sensors and alarms monitor the environment surrounding the equipment to ensure that air and cooling water temperatures remain within the levels specified by equipment design. If proper conditions are not maintained, the alarms summon operations and maintenance personnel to correct the situation before a business interruption occurs.

3.2 Technical Controls

Technical security involves the use of safeguards incorporated in computer hardware, operations or applications software, communications hardware and software, and related devices. Technical controls are sometimes referred to as logical controls.

Preventive Technical Controls
Preventive technical controls are used to prevent unauthorized personnel or programs from gaining remote access to computing resources. Examples of these controls include:

- Access control software.
- Antivirus software.
- Library control systems.
- Passwords.
- Smart cards.
- Encryption.
- Dial-up access control and callback systems.

Access Control Software: The purpose of access control software is to control sharing of data and programs between users. In many computer systems, access to data and programs is implemented by access control lists that designate which users are allowed access. Access control software provides the ability to control access to the system by establishing that only registered users with an authorized log-on ID and password can gain access to the computer system.

After access to the system has been granted, the next step is to control access to the data and programs residing in the system. The data or program owner can establish rules that designate who are authorized to use the data or program.

Antivirus Software: Viruses have reached epidemic proportions throughout the micro computing world and can cause processing disruptions and loss of data as well as significant loss of productivity while cleanup is conducted. In addition, new viruses are emerging at an ever-increasing rate-currently about one every 48 hours. It is

recommended that antivirus software be installed on all microcomputers to detect, identify, isolate, and eradicate viruses. This software must be updated frequently to help fight new viruses. In addition, to help ensure that viruses are intercepted as early as possible, antivirus software should be kept active on a system, not used intermittently at the discretion of users.

Library Control Systems: These systems require that all changes to production programs be implemented by library control personnel instead of the programmers who created the changes. This practice ensures separation of duties, which helps prevent unauthorized changes to production programs.

Passwords

Passwords are used to verify that the user of an ID is the owner of the ID. The ID-password combination is unique to each user and therefore provides a means of holding users accountable for their activity on the system.

Fixed passwords that are used for a defined period of time are often easy for hackers to compromise; therefore, great care must be exercised to ensure that these passwords do not appear in any dictionary. Fixed passwords are often used to control access to specific data bases. In this use, however, all persons who have authorized access to the data base use the same password; therefore, no accountability can be achieved. Currently, dynamic or one-time passwords, which are different for each log-on, are preferred over fixed passwords. Dynamic passwords are created by a token that is programmed to generate passwords randomly.

Smart Cards: Smart cards are usually about the size of a credit card and contain a chip with logic functions and information that can be read at a remote terminal to identify a specific user's privileges. Smart cards now carry prerecorded, usually encrypted access control information that is compared with data that the user provides (e.g., a personal ID number or biometric data) to verify authorization to access the computer or network.

Encryption: Encryption is defined as the transformation of plaintext (i.e., readable data) into ciphertext (i.e., unreadable data) by cryptographic techniques. Encryption is currently considered to be the only sure way of protecting data from disclosure during network transmissions.

Encryption can be implemented with either hardware or software. Software-based encryption is the least expensive method and is suitable

for applications involving low-volume transmissions; the use of software for large volumes of data results in an unacceptable increase in processing costs. Because there is no overhead associated with hardware encryption, this method is preferred when large volumes of data are involved.

Dial-Up Access Control and Callback Systems: Dial-up access to a computer system increases the risk of intrusion by hackers. In networks that contain personal computers or are connected to other networks, it is difficult to determine whether dial-up access is available or not because of the ease with which a modem can be added to a personal computer to turn it into a dial-up access point. Known dial-up access points should be controlled so that only authorized dial-up users can get through.

Currently, the best dial-up access controls use a microcomputer to intercept calls, verify the identity of the caller (using a dynamic password mechanism), and switch the user to authorized computing resources as requested. Previously, call-back systems intercepted dial-up callers, verified their authorization and called them back at their registered number, which at first proved effective; however, sophisticated hackers have learned how to defeat this control using call-forwarding techniques.

Detective Technical Controls

Detective technical controls warn personnel of violations or attempted violations of preventive technical controls. Examples of these include audit trails and intrusion detection expert systems, which are discussed in the following sections.

Audit Trails: An audit trail is a record of system activities that enables the reconstruction and examination of the sequence of events of a transaction, from its inception to output of final results. Violation reports present significant, security-oriented events that may indicate either actual or attempted policy transgressions reflected in the audit trail. Violation reports should be frequently and regularly reviewed by security officers and data base owners to identify and investigate successful or unsuccessful unauthorized accesses.

Intrusion Detection Systems: These expert systems track users (on the basis of their personal profiles) while they are using the system to determine whether their current activities are consistent with an established norm. If not, the user's session can be terminated or a security officer can be called to investigate. Intrusion detection can be especially effective in cases in which intruders are pretending to be authorized users or when authorized users are involved in unauthorized activities.

3.3 Administrative Controls

Administrative, or personnel, security consists of management constraints, operational procedures, accountability procedures, and supplemental administrative controls established to provide an acceptable level of protection for computing resources. In addition, administrative controls include procedures established to ensure that all personnel who have access to computing resources have the required authorizations and appropriate security clearances.

Preventive Administrative Controls
Preventive administrative controls are personnel-oriented techniques for controlling people's behavior to ensure the confidentiality, integrity, and availability of computing data and programs. Examples of preventive administrative controls include:

- Security awareness and technical training.
- Separation of duties.
- Procedures for recruiting and terminating employees.
- Security policies and procedures.
- Supervision.
- Disaster recovery, contingency, and emergency plans.
- User registration for computer access.

Security Awareness and Technical Training: Security awareness training is a preventive measure that helps users to understand the benefits of security practices. If employees do not understand the need for the controls being imposed, they may eventually circumvent them and thereby weaken the security program or render it ineffective.

Technical training can help users prevent the most common security problem—errors and omissions—as well as ensure that they understand how to make appropriate backup files and detect and control viruses. Technical training in the form of emergency and fire drills for operations personnel can ensure that proper action will be taken to prevent such events from escalating into disasters.

Separation of Duties: This administrative control separates a process into component parts, with different users responsible for different parts of the process. Judicious separation of duties prevents one individual from obtaining control of an entire process and forces collusion with others in order to manipulate the process for personal gain.

Recruitment and Termination Procedures: Appropriate recruitment procedures can prevent the hiring of people who are likely to violate

security policies. A thorough background investigation should be conducted, including checking on the applicant's criminal history and references. Although this does not necessarily screen individuals for honesty and integrity, it can help identify areas that should be investigated further.

Three types of references should be obtained: (1) employment, (2) character, and (3) credit. Employment references can help estimate an individual's competence to perform, or be trained to perform, the tasks required on the job. Character references can help determine such qualities as trustworthiness, reliability, and ability to get along with others. Credit references can indicate a person's financial habits, which in turn can be an indication of maturity and willingness to assume responsibility for one's own actions.

In addition, certain procedures should be followed when any employee leaves the company, regardless of the conditions of termination. Any employee being involuntarily terminated should be asked to leave the premises immediately upon notification, to prevent further access to computing resources. Voluntary terminations may be handled differently, depending on the judgment of the employee's supervisors, to enable the employee to complete work in process or train a replacement.

All authorizations that have been granted to an employee should be revoked upon departure. If the departing employee has the authority to grant authorizations to others, these other authorizations should also be reviewed. All keys, badges, and other devices used to gain access to premises, information, or equipment should be retrieved from the departing employee. The combinations of all locks known to a departing employee should be changed immediately. In addition, the employee's log-on IDs and passwords should be canceled, and the related active and backup files should be either deleted or reassigned to a replacement employee.

Any special conditions to the termination (e.g., denial of the right to use certain information) should be reviewed with the departing employee; in addition, a document stating these conditions should be signed by the employee. All terminations should be routed through the computer security representative for the facility where the terminated employee works to ensure that all information system access authority has been revoked.

Security Policies and Procedures: Appropriate policies and procedures are key to the establishment of an effective information security program. Policies and procedures should reflect the general policies of the organization as regards the protection of information and computing

resources. Policies should cover the use of computing resources, marking of sensitive information, movement of computing resources outside the facility, introduction of personal computing equipment and media into the facility, disposal of sensitive waste, and computer and data security incident reporting. Enforcement of these policies is essential to their effectiveness.

Supervision: Often, an alert supervisor is the first person to notice a change in an employee's attitude. Early signs of job dissatisfaction or personal distress should prompt supervisors to consider subtly moving the employee out of a critical or sensitive position.

Supervisors must be thoroughly familiar with the policies and procedures related to the responsibilities of their department. Supervisors should require that their staff members comply with pertinent policies and procedures and should observe the effectiveness of these guidelines. If the objectives of the policies and procedures can be accomplished more effectively, the supervisor should recommend appropriate improvements. Job assignments should be reviewed regularly to ensure that an appropriate separation of duties is maintained, that employees in sensitive positions are occasionally removed from a complete processing cycle without prior announcement, and that critical or sensitive jobs are rotated periodically among qualified personnel.

Disaster Recovery, Contingency, and Emergency Plans: The disaster recovery plan is a document containing procedures for emergency response, extended backup operations, and recovery should a computer installation experience a partial or total loss of computing resources or physical facilities (or of access to such facilities). The primary objective of this plan, used in conjunction with the contingency plans, is to provide reasonable assurance that a computing installation can recover from disasters, continue to process critical applications in a degraded mode, and return to a normal mode of operation within a reasonable time. A key part of disaster recovery planning is to provide for processing at an alternative site during the time that the original facility is unavailable.

Contingency and emergency plans establish recovery procedures that address specific threats. These plans help prevent minor incidents from escalating into disasters. For example, a contingency plan might provide a set of procedures that defines the condition and response required to return a computing capability to nominal operation; an emergency plan might be a specific procedure for shutting down equipment in the event of a fire or for evacuating a facility in the event of an earthquake.

User Registration for Computer Access: Formal user registration ensures that all users are properly authorized for system and service access. In addition, it provides the opportunity to acquaint users with their responsibilities for the security of computing resources and to obtain their agreement to comply with related policies and procedures.

Detective Administrative Controls
Detective administrative controls are used to determine how well security policies and procedures are complied with, to detect fraud, and to avoid employing persons that represent an unacceptable security risk. This type of control includes:

- Security reviews and audits.
- Performance evaluations.
- Required vacations.
- Background investigations.
- Rotation of duties.

Security Reviews and Audits: Reviews and audits can identify instances in which policies and procedures are not being followed satisfactorily. Management involvement in correcting deficiencies can be a significant factor in obtaining user support for the computer security program.

Performance Evaluations: Regularly conducted performance evaluations are an important element in encouraging quality performance. In addition, they can be an effective forum for reinforcing management's support of information security principles.

Required Vacations: Tense employees are more likely to have accidents or make errors and omissions while performing their duties. Vacations contribute to the health of employees by relieving the tensions and anxieties that typically develop from long periods of work. In addition, if all employees in critical or sensitive positions are forced to take vacations, there will be less opportunity for an employee to set up a fraudulent scheme that depends on the employee's presence (e.g., to maintain the fraud's continuity or secrecy). Even if the employee's presence is not necessary to the scheme, required vacations can be a deterrent to embezzlement because the employee may fear discovery during his or her absence.

Background Investigations: Background investigations may disclose past performances that might indicate the potential risks of future performance. Background investigations should be conducted on all employees being considered for promotion or transfer into a position of trust; such investigations should be completed before the employee is

actually placed in a sensitive position. Job applicants being considered for sensitive positions should also be investigated for potential problems. Companies involved in government-classified projects should conduct these investigations while obtaining the required security clearance for the employee.

Rotation of Duties: Like required vacations, rotation of duties (i.e., moving employees from one job to another at random intervals) helps deter fraud. An additional benefit is that as a result of rotating duties, employees are cross-trained to perform each other's functions in case of illness, vacation, or termination.

4.0 CONCLUSION

Information security controls can be classified as physical, technical, or administrative. These are further divided into preventive and detective controls.

The organization's security policy should be reviewed to determine the confidentiality, integrity, and availability needs of the organization. The appropriate physical, technical, and administrative controls can then be selected to provide the required level of information protection, as stated in the security policy.

A careful balance between preventive and detective control measures is needed to ensure that users consider the security controls reasonable and to ensure that the controls do not overly inhibit productivity. The combination of physical, technical, and administrative controls best suited for a specific computing environment can be identified by completing a quantitative risk analysis. Because this is usually an expensive, tedious, and subjective process, however, an alternative approach-referred to as meeting the standard of due care-is often used. Controls that meet a standard of due care are those that would be considered prudent by most organizations in similar circumstances or environments.

5.0 SUMMARY

- Security is generally defined as the freedom from danger or as the condition of safety. Computer security, specifically, is the protection of data in a system against unauthorized disclosure, modification, or destruction and protection of the computer system itself against unauthorized use, modification, or denial of service.

- Physical security is the use of locks, security guards, badges, alarms, and similar measures to control access to computers, related equipment (including utilities), and the processing facility itself.
- An audit trail is a record of system activities that enables the reconstruction and examination of the sequence of events of a transaction, from its inception to output of final results
- Administrative, or personnel, security consists of management constraints, operational procedures, accountability procedures, and supplemental administrative controls established to provide an acceptable level of protection for computing resources.
- Detective administrative controls are used to determine how well security policies and procedures are complied with, to detect fraud, and to avoid employing persons that represent an unacceptable security risk.

6.0 TUTOR-MARKED ASSIGNMENT

1. Identify 5 components of Detective Administrative Control.
2. Briefly discuss password as a technical control against illegal access to data.

7.0 REFERENCES/FURTHER READINGS

Krause, M. and Tipton, H.F. Handbook of Information Security *Management.*

UNIT 5 NETWORK SECURITY AND MANAGEMENT

CONTENTS

1.0 INTRODUCTION

Network Management as a term has many definitions dependent on whose operational function is in question. It is the goal of this unit to illustrate and discuss today's most common implementations of network management systems as they apply to actual MIS form and function and illustrate what's wrong with this picture type of scenario. Also, the unit will discuss what the ideal system will look like.

Network security consists of the provisions made in an underlying network infrastructure, policies adopted by the network administrator to protect the network and the network-accessible resources from unauthorized access and the effectiveness (or lack) of these measures combined together.

Network management systems have been in operation many years especially in their own proprietary worlds such as Netview, AT&T Accumaster and Digital Equipment Corporation's DMA. With the implementation of SNMP, local area and wide area network components could be monitored and "managed". With the vast amount of raw data available, most MIS Managers have no idea what they really want because, in part, they don't know what's available. Additionally, how does the data get into a format that actually means something? Other communications systems are considered non-manageable because they are only accessible by an RS-232 port and not by Netview or SNMP. Others tend to believe that Network Management means nothing but the monitoring and management of network architectural hardware such as

Routers, bridges and concentrators -- nothing above the network layer of the OSI model is considered manageable.

Network management doesn't mean one application with a database with some huge chunk of iron running the show. It is really an integrated conglomeration of functions that may be on one machine but may span thousands of miles, different support organizations and many machines and databases. It is these functions that must be directly driven by the business case for each.

2.0 OBJECTIVES

At the end of this unit you should be able to:

- define network security and management
- identify the attributes of a secure network
- identify the components of a functional architecture of network management
- explain the levels of network management and how they differ from one another
- understand and identify the constituents of Management Functional Areas (MFAs)
- define how to implement network security and management.

3.0 MAIN CONTENT

3.1 Attributes of a Secure Network

Network security starts from authenticating any user, most likely a username and a password. Once authenticated, a stateful firewall enforces access policies such as what services are allowed to be accessed by the network users. Though effective to prevent unauthorized access, this component fails to check potentially harmful contents such as computer worms being transmitted over the network. An intrusion prevention system (IPS)[2] helps detect and prevent such malware. IPS also monitors for suspicious network traffic for contents, volume and anomalies to protect the network from attacks such as denial of service. Communication between two hosts using the network could be encrypted to maintain privacy. Individual events occurring on the network could be tracked for audit purposes and for a later high level analysis.

Honeypots, essentially decoy network-accessible resources, could be deployed in a network as surveillance and early-warning tools. Techniques used by the attackers that attempt to compromise these decoy resources are studied during and after an attack to keep an eye on

new exploitation techniques. Such analysis could be used to tighten security of the actual network being protected by the honeypot.

3.2 Functional Architecture

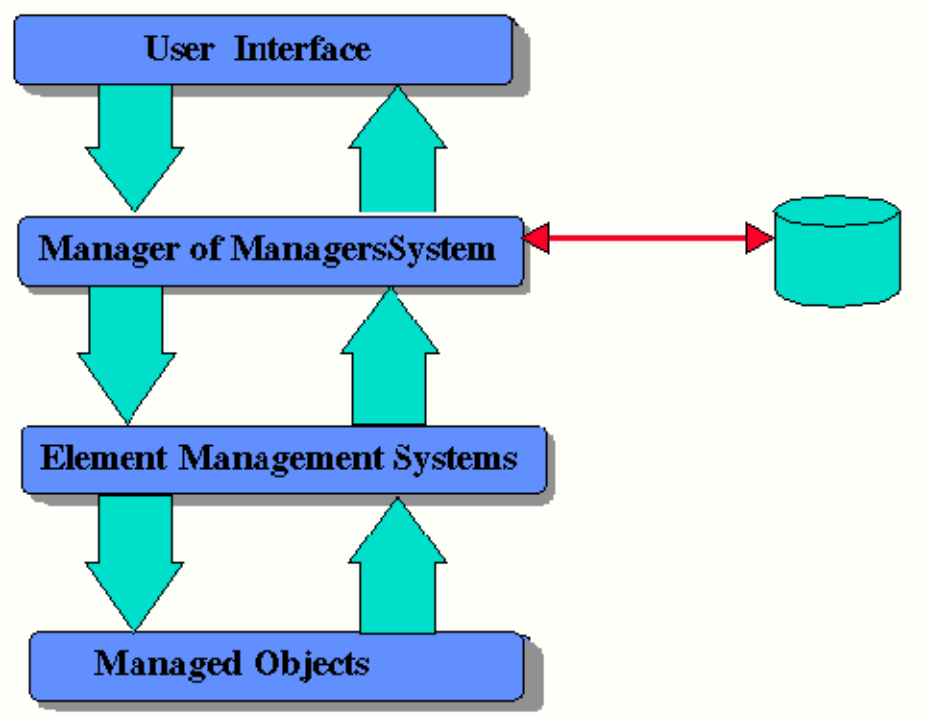


Figure 1

Defining the Pieces

Network management systems have four basic levels of functionality. Each level has a set of tasks defined to provide, format, or collect data necessary to manage the objects. Figure 1 illustrates these four levels of functionality.

Managed Objects

Managed Objects are the devices, systems and/or anything else requiring some form of monitoring and management. Most implementations leave out the "anything else" clause because they usually don't have the business case requirements before the design, therefore they design as they go.

Some examples of managed objects include routers, concentrators, hosts, servers and applications like Oracle, Microsoft SMS, Lotus Notes, and MS Mail. The managed object does not have to be a piece of

hardware but should rather be depicted as a function provided on the network.

Element Management Systems (EMS)

An EMS manages a specific portion of the network. For example SunNet Manager, an SNMP management application, is used to manage SNMP manageable elements. Element Managers may manage async lines, multiplexers, PABX's, proprietary systems or an application.

Manager of Managers Systems (MoM)

MoM systems integrate together the information associated with several element management systems, usually performing alarm correlation between EMS's. There are several different products that fall into this category to include Boole & Babbage's CommandPost, NyNEX AllLink, International Telematics MAXM, OSI NetExpert and others.

The actual data to be collected comes from the managed object, in most cases. This data is collected by the EMS systems which in turn consolidates the data in a database for processing and retrieval.

User Interface

The user interface to the information, whether real time alarms and alerts or trend analysis graphs and reports, is the principal piece to deploying a successful system. If the information gathered cannot be distributed to the whole MIS organization to keep people informed and to enable team communications, the real purpose of a Network Management system is lost in the implementation. Data doesn't mean anything if it is not used to make informed decisions about the optimization of systems and functions.

These systems components are, in turn, mapped back to what is called Management Functional Areas (MFAs). These MFAs are the wish list of which areas in which management applications as a system focus their attention.

3.3 Levels of Security Management

Security Management for networks is different for all kinds of situations. A small home or an office would only require basic security while large businesses will require high maintenance and advanced software and hardware to prevent malicious attacks from hacking and spamming.

Small Homes

- A basic firewall
- For Windows users, basic Antivirus software like McAfee, Norton AntiVirus, AVG Antivirus or Windows Defender, others may suffice if they contain a virus scanner to scan for malicious software.
- When using a wireless connection, use a robust password.

Medium Businesses

- A fairly strong firewall
- A strong Antivirus software and Internet Security Software.
- For authentication, use strong passwords and change it on a bi- weekly/monthly basis.
- When using a wireless connection, use a robust password.
- Raise awareness about physical security to employees.
- Use an optional network analyzer or network monitor.

Large Businesses

- A strong firewall and proxy to keep unwanted people out.
- A strong Antivirus software and Internet Security Software.
- For authentication, use strong passwords and change it on a weekly/bi-weekly basis.
- When using a wireless connection, use a robust password.
- Exercise physical security precautions to employees.
- Prepare a network analyzer or network monitor and use it when needed.
- Implement physical security management like closed circuit television for entry areas and restricted zones.
- Security fencing to mark the company's perimeter.
- Fire extinguishers for fire-sensitive areas like server rooms and security rooms.
- Security guards can help to maximize security.

School

- An adjustable firewall and proxy to allow authorized users access from the outside and inside.
- A strong Antivirus software and Internet Security Software.
- Wireless connections that lead to firewalls.
- CIPA compliance.
- Supervision of network to guarantee updates and changes based on popular site usage.
- Constant supervision by teachers, librarians, and administrators to guarantee protection against attacks by both Internet and sneakernet

sources.

Large Government

- A strong strong firewall and proxy to keep unwanted people out.
- A strong Antivirus software and Internet Security Software.
- Strong encryption, usually with a 256 bit key.
- Whitelist authorized wireless connection, block all else.
- All network hardware is in secure zones.
- All hosts should be on a private network that is invisible from the outside.
- Put all servers in a DMZ, or a firewall from the outside and from the inside.
- Security fencing to mark perimeter and set wireless range to this.

3.4 Management Functional Areas (MFAs)

The most common framework depicted in Network management designs is centered around the Open Systems Interconnect (OSI) "FCAPS" model of MFAs. However most network management implementations do not really cover all of these areas. Other areas that may be important to the e-business/MIS function and to specific business units within the company may not be addressed at all.

FCAPS is an acronym explained as follows:

- Fault Management
- Configuration Management
- Accounting
- Performance Management
- Security Management

Some of the other areas covered under Management Functional Areas include:

- Chargeback
- Systems Management
- Cost Management

- Fault Management

Fault management is the detection of a problem, fault isolation and correction to normal operation. Most systems poll the managed objects search for error conditions and illustrate the problem in either a graphic format or a textual message. Most of these types of messages are setup

by the person configuring the polling on the Element Manager. Some Element Management Systems collect data directly from a log printer type output receiving the alarm as it occurs.

Fault management deals most commonly with events and traps as they occur on the network. Keep in mind though, that using data reporting mechanisms to report alarms or alerts is the best way to accomplish health checks of specific managed object's performance without having to double the amount of polling being accomplished.

• Configuration Management

Configuration management is probably, the most important part of network management in that you cannot accurately manage a network unless you can manage the configuration of the network. Changes, additions and deletions from the network need to be coordinated with the network management systems personnel. Dynamic updating of the configuration needs to be accomplished periodically to ensure the configuration is known.

• Accounting

The accounting function is usually left out of most implementations in that LAN based systems are said to not promote accounting type functions until one gets into the Hosts such as IBM Mainframe or Digital VAX's. Others rationalize the accounting is a server specific function and should be managed by the System administrators.

• Performance Management

Performance is a key concern to most MIS support people. Although, it is high on the list, it is considered difficult to be factual about some LAN performance issues unless employing RMON technology. (This is one of those examples of throwing money at a problem.) Although Pods are very useful, one should carefully weigh what's pertinent to what can be accomplished in other ways without having to spend a bundle.

Performance of Wide Area Network (WAN) links, telephone trunk utilization, etc., are areas that must be revisited on a continuing basis as these are some of the areas easiest to optimize and realize savings.

Systems or applications performance is another area in which optimization can be accomplished but most network management applications don't address this in a functional manner.

• Security Management

Most network management applications only address security applicable to network hardware such as someone logging into a router or bridge. Some network management systems have alarm detection and reporting capabilities as part of physical security (contact closure, fire alarm interface, etc.) None really deal with system security as this is a function of System administration.

Chargeback

Chargeback has been done for years in the large mainframe environments and will continue to be accomplished as it is a way to charge the end user for only the specific portion of the service that he or she uses. Chargeback on Local Area Networks presents new challenges in that so many services are provided. In many implementations, chargeback is accomplished on the individual Server providing the service. While chargeback is very difficult on broadcast based networks such as Ethernet, it is realizable on networks that dynamically allocate bandwidth as the end users' needs dictate (ATM). As technology associated with monitoring LAN and WAN networks evolves, chargeback will be integrated into more and more systems.

• Systems Management

Systems Management is the management and administration of services provided on the network. A lot of implementations leave out this very crucial part in that this is one of the areas in which Network management systems can show significant capabilities, streamline business processes, and save the customer money with just a little work. There are many good COTS products available to automate system administration functions and these products can be easily integrated into the overall Network Management system very easily.

• Cost Management

Cost management is an avenue in which the reliability, operability and maintainability of managed objects are addressed. This one function is an enabler to upgrade equipment, delete unused services and tune the functionality of the Servers to the services provided. By continuously addressing the cost of maintenance, Mean Time Between Failure (MTBF), and Mean Time To Repair (MTTR) statistics, costs associated with maintaining the network as a system can be tuned. This area is an MFA that is driven by I/T management to address getting the most performance from the money allocated.

3.5 Common Implementations

Most implementations of medium and large network management systems center on a Network Management Center of some sort. From this location, all data is sent and processed. While several EMS's are used to manage their specific areas, all of the data comes back to the Manager of Managers application. Most fault detection, isolation and troubleshooting is accomplished in the Network Management Center and technicians dispatched when the problem has been analyzed as far as possible. Several company locations may be involved in the overall network spanning thousands of miles and around the globe.

Management Focus

The management focus for this scenario is on the Network Management Center driving the total operation. Detection, troubleshooting and dispatching is accomplished from the NMC. This operational focus is a carry over from the old Netview days in that the center of the picture was a huge IBM Mainframe that did all of the work. If you don't have a Network Management Center today, consider what it will cost not only for the hardware and software, but the people to accomplish this and their level of expertise.

The Right Implementation

If you, as an MIS Manager, are looking at the benefits of network management to reduce downtime and overall cost to your company, make sure that the business case requirements drive the implementation and not the implementation drive the business cases.

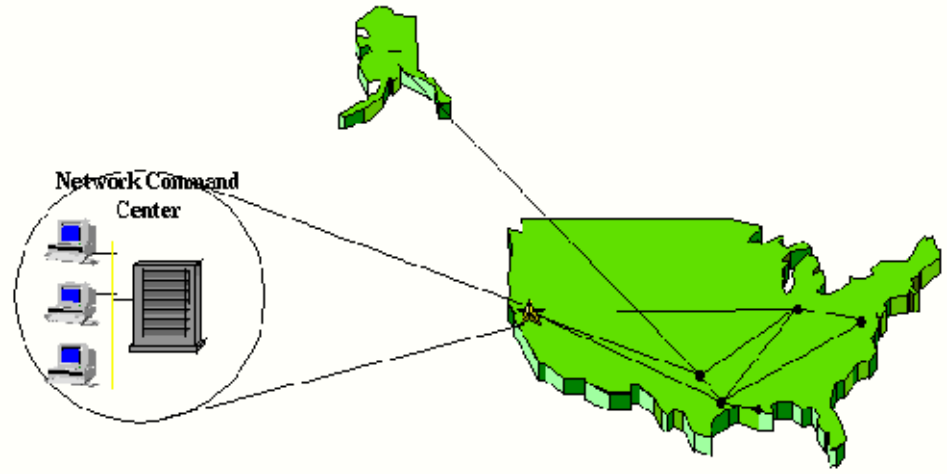


Figure 2

As a systems integrator, make sure the requirements are accomplished before any implementation. When the requirements are put in place, it is

your job as an Engineer to make sure management is informed as to what each implementation segment will cost along with what that capability brings to the overall MIS function.

3.6 Business/E-Business Case requirements

In today's world, any implementation must follow the business case associated with what will be implemented. The implementation must solve a business problem or increase efficiency of the current methods of accomplishing work while reducing overall costs. If the solution does not save money while providing a better service, it probably is not worth accomplishing.

Definition

The hardest part of building a business case is the gathering of the information. One must define the problem at hand in a general sense so that you can look for specific problems network management can address in that area.

The developer of the business case must look at the current way each section accomplishes its day to day work. The case for network management can be definitised by documenting current work processes that may be automated by the system as a whole. Each of the work processes to be automated need to be documented and addressed in the system design and implementation.

Look for ways to save the organization money. Keep addressing getting the MIS organization and the services they provide, more efficient.

Levels of Activity

There are four levels of activity that one must understand before applying management to a specific service or device. These four levels of activity are as follows:

• Inactive

This is the case when no monitoring is being done and if you did receive an alarm in this area, you would ignore it.

• Reactive

This is where you react to a problem after it has occurred yet no monitoring has been applied.

• Interactive

This is where you are monitoring components but must interactively troubleshoot to eliminate the side effect alarms and isolate to a root cause.

• Proactive

This is where you are monitoring components and the system provides a root cause alarm for the problem at hand and automatic processes are in place where possible to minimize downtime.

These four levels of activities outline exactly how your support organization is dealing with problems today and where you, as an MIS manager want them to be in terms of goals. Within the organization are teams with different goals and focuses (i.e. Unix support, desktop support, network support, etc.). Keep in mind while a specific alarm may warrant an inactive approach by one team, to another team it may demand a proactive approach. Keep these goals in mind when gathering requirements for network management.

Today's Implementations

Of the network management implementations done today, very few really address the needs of the business. Most are implemented with good intentions but are focused away from increasing efficiency.

In a multiple site network, there are technicians, engineers and support personnel at each major location as required. No one knows those local environments better than the people having to do the work. No one knows the people of the organization better than the Help Desk staff as they are the first line of communication between the people and the MIS support organization.

Network management elements are considered, among other things, tools in which troubleshooting can be accomplished. The local support staff could benefit greatly from the use of these systems as a tool. As such, most implementations give read-only access to these systems. The ability to focus these tools at a local level is paramount to increasing the effectiveness to the local support staff. In some implementations, where read/write access is provided, it is accomplished through X-Windows which doesn't work very well across low speed links.

Most implementations focus these tools at a global level in that they are located in the Network Command Center. When a trouble ticket is

generated from the NCC, it reflects a problem or symptom generated by the network management elements and/or the Manager of Managers. Sometimes, the local technician can not relate to this symptom because he or she doesn't understand where this message came from or why. Without access to the management element and familiarity with the product, they usually start off problem isolation in a "cloud" looking for the problem.

When a global problem occurs, in these scenarios, the information is concentrated and orchestrated by the Network Command Center. Additionally an outage can black out management of a geographic location by centralizing the management resources. Figure 3 illustrates how this occurs.

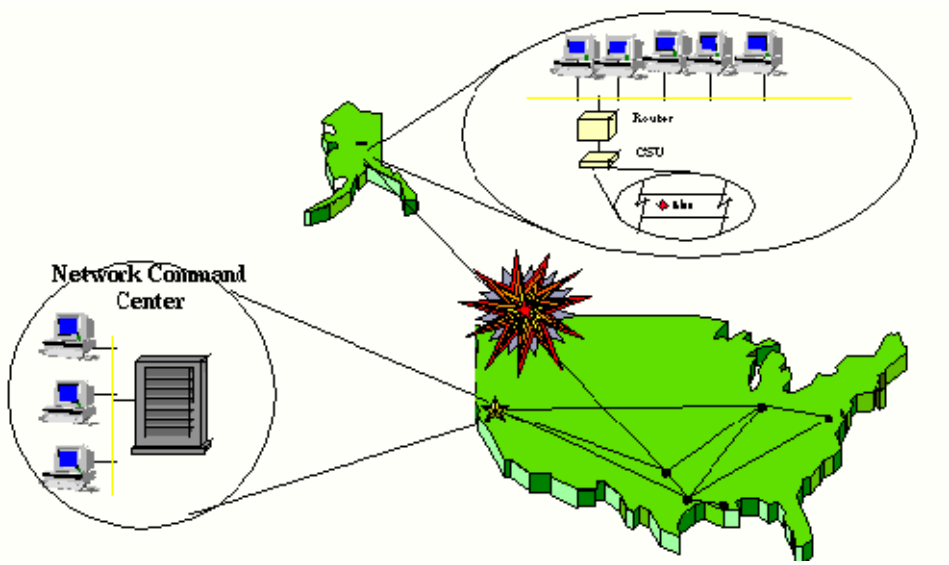


Figure 3

As far as the Network Management Center is concerned, all of the devices beyond the point of breakage are down. In fact, without alarm correlation, all of the devices will be depicted as bad. Even with alarm correlation, it can only be accomplished on one side of the link. No network management capabilities exist at the remote site to help troubleshoot the problem.

System Focus

The ideal network management system should be designed and implemented around the real work processes. It should focus the tools toward those staff members supporting the managed area in a manner which makes their job easier and faster. Information associated with a problem or symptom should mean something to the support personnel. If they see the problem at a glance, they should know which specific

area that problem belongs and what to do to get started in the trouble isolation process. Other personnel in the organization should know that a specific technician is looking into the problem as the problem may be affecting other areas.

Help Desk personnel should know what is happening and who is working on what at a glance. If they are not familiar with the system in question, they should have adequate information at their fingertips to guide them in what to do, who to call, and what steps to take, even what questions to ask.

Additionally, the problems that affect other sites, should be available to those personnel at a glance. The information must be at the fingertips of the other sites' Help Desk personnel so that they know, in near real time, what is going on.

See how the focus of information should be; local when it is a local problem and global when it is a global problem. Also, the tools associated are more focused on the local situation and not the global picture.

Figure 4 depicts a more distributed system providing global information with local focus. In this system, alarms can be passed from site to site and even around a problem with simple client-server database techniques.

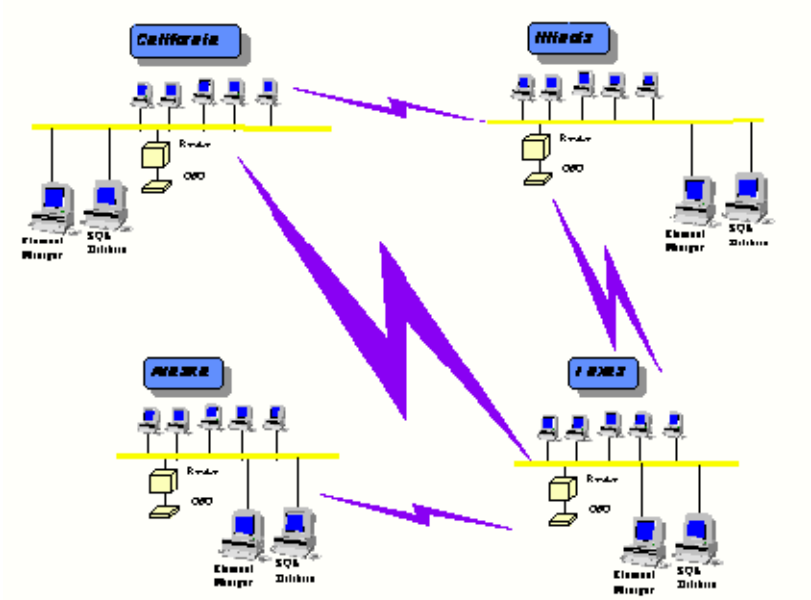


Figure 4

In the scenario in figure 4, if a link breaks, local tools and alarms are still available. Alarms concerning the overall health of other links and connectivity can be passed to other sites, even around a problem. Using

a SLIP or PPP dial up link between management elements can be used to pass critical data about a link outage in near real time.

Network management across low speed wide area links doesn't really make sense. Bandwidth of this type is costly compared to LAN bandwidth in that there are the monthly charges for the links. Consider also that most WAN links are interconnected by bridges or routers. On the back side of these devices are networks capable of 10 Mbps, 16 Mbps or even 100 Mbps. On the link side you see 1.544 Mbps, 512kbps or even 19.2kbps links. Actual polling of network management elements (SNMP) could consume these links drastically reducing the operational capabilities of the link. The question to ask is Do you want to increase *the bandwidth across these links just for network management or do you want to distribute the management polling to local area concentrations and just pass the real alarm information?*

4.0 CONCLUSION

There are a lot of excellent products available today that provide capabilities to manage not just hardware, but services and applications. The way that these systems are implemented are also critical in that each management capability installed must match a business need for such a system. Additionally, these diverse systems must be integrated together and into the support organizations to achieve maximum effectiveness.

5.0 SUMMARY

- Network security consists of the provisions made in an underlying computer network infrastructure, policies adopted by the network administrator to protect the network and the network-accessible resources from unauthorized access and the effectiveness (or lack) of these measures combined together.
- Network security starts from authenticating any user, most likely a username and a password. Once authenticated, a stateful firewall enforces access policies such as what services are allowed to be accessed by the network users.
- Network management systems have four basic levels of functionality. Each level has a set of tasks defined to provide, format, or collect data necessary to manage the objects.
- Security Management for networks is different for all kinds of situations. A small home or an office would only require basic security while large businesses will require high maintenance and advanced software and hardware to prevent malicious attacks from hacking and spamming.

- The most common framework depicted in Network management designs is centered around the Open Systems Interconnect (OSI) "FCAPS" model of MFAs.
- Most implementations of medium and large network management systems center around a Network Management Center of some sort. From this location, all data is sent and processed. While several EMS's are used to manage their specific areas, all of the data comes back to the Manager of Managers application.
- In today's world, any implementation must follow the business case associated with what will be implemented. The implementation must solve a business problem or increase efficiency of the current methods of accomplishing work while reducing overall costs.

6.0 TUTOR-MARKED ASSIGNMENT

1. List 5 security management for a medium-scale business.
2. Briefly discuss the Levels of Activity in Business/e-business Case Requirements.

7.0 REFERENCES/FURTHER READINGS

Douglas W. Stevenson, (1995). itmWEB Media Corporation.

A Role-Based Trusted Network Provides Pervasive Security and Compliance - interview with Jayshree Ullal, Senior VP of Cisco

Dave Dittrich, Network monitoring/Intrusion Detection Systems (IDS). University of Washington.

Honeypots, Honeynets

Security of the Internet (The Froehlich/Kent Encyclopedia of *Telecommunications vol. 15. Marcel Dekker, New York, 1997,* pp. 231-255.)

Introduction to Network Security, Matt Curtin.

Security Monitoring with Cisco Security MARS, Gary Halleen/Greg Kellogg, Cisco Press, Jul. 6, 2007.

Self-Defending Networks: The Next Generation of Network Security, Duane DeCapite, Cisco Press, Sep. 8, 2006.

Security Threat Mitigation and Response: Understanding CS-MARS, Dale Tesch/Greg Abelar, Cisco Press, Sep. 26, 2006.

Deploying Zone-Based Firewalls, Ivan Pepelnjak, Cisco Press, Oct. 5, 2006.

Network Security: PRIVATE Communication in a PUBLIC World, Charlie Kaufman | Radia Perlman | Mike Speciner, Prentice-Hall, 2002.

MODULE 2

Unit 1 Business Continuity Planning
Unit 2 Information Systems Security Controls
Unit 3 Information Systems Audit
Unit 4 Copyright Law and Electronic Access to Information
Unit 5 Internet Firewall

UNIT 1 BUSINESS CONTINUITY PLANNING

CONTENTS

1.0 INTRODUCTION

Today's organisations, in their efforts to reduce costs, are streamlining layers of management while implementing more complex matrices of control and reporting. Distributed systems have facilitated the reshaping of these organisations by moving the control of information closer to its source, the end user. In this transition, however, secure management of that information has been placed at risk. Information technology departments must protect the traditional system environment within the computer room plus develop policies, standards, and guidelines for the security and protection of the company's distributed information base. Further, the information technology staff must communicate these standards to all users to enforce a strong baseline of controls.

In these distributed environments, information technology personnel are often asked to develop systems recovery plans outside the context of an overall business recovery scheme. Recoverability of systems, however, should be viewed as only one part of business recovery. Information systems, in and of themselves, are not the lifeblood of a company; inventory, assets, processes, and people are all essential factors that must be considered in the business continuation design. The success of

business continuity planning rests on a company's ability to integrate systems recovery in the greater overall planning effort.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the success factor for business continuity
- explain the phases and how to plan for business recovery
- understand how to appraise management of risks and mitigation costs
- understand how to plan for distributed environment
- discuss the network recovery strategies available.

3.0 MAIN CONTENT

3.1 Business Recovery Planning-The Process

Distinctive areas must be addressed in the formulation of a company's disaster recovery plan, and attention to these areas should follow the steps of the scientific method: a statement of the problem, the development of a hypothesis, and the testing of the hypothesis. Like any scientific process, the development of the disaster recovery plan is iterative. The testing phase of this process is essential because it reveals whether the plan is viable. Moreover, it is imperative that the plan and its assumptions be tested on an ongoing, routine basis. The most important distinction that marks disaster recovery planning is what is at stake—the survival of the business.

The phases of a disaster recovery plan process are:

- Awareness and discovery
- Risk assessment
- Mitigation
- Preparation
- Testing
- Response and recovery

Recovery planners should adapt these phases to a company's specific needs and requirements. Some of the phases may be combined, for example, depending on the size of the company and the extent of exposures to risk. It is crucial, however, that each phase be included in the formation of a recovery plan.

• Awareness and Discovery

Awareness begins when a recovery planning team can identify both possible threats and plausible threats to business operations. The more pressing issue for an organisation in terms of business recovery planning is that of plausible threats. These threats must be evaluated by recovery planners, and their planning efforts, in turn, will depend on these criteria:

• The business of the company.
• The area of the country in which the company is located.
• The company's existing security measures.
• The level of adherence to existing policies and procedures.
• Management's commitment to existing policies and procedures.

Awareness also implies educating all employees on existing risk exposures and briefing them on what measures have been taken to minimize those exposures. Each employee's individual role in complying with these measures should be addressed at this early stage.

In terms of systems and information, the awareness phase includes determining what exposures exist that are specific to information systems, what information is vital to the organisation, and what information is proprietary and confidential. Answering these questions will help planners determine when an interruption will be catastrophic as opposed to operational. For example, in an educational environment, a system that is down for two or three days may not be catastrophic, whereas in a process control environment (e.g., chemicals or electronics), just a few minutes of downtime may be.

Discovery is the process in which planners must determine, based on their awareness of plausible threats, which specific operations would be affected by existing exposures. They must consider what measures are currently in place or could be put in place to minimize or, ideally, remove these exposures.

• Risk Assessment

Risk assessment is a decision process that weighs the cost of implementing preventive measures against the risk of loss from not implementing them. There are many qualitative and quantitative approaches to risk analysis. Typically, two major cost factors arise for the systems environment: the first is the loss incurred from a cease in business operations due to system downtime, and the second is the replacement cost of equipment.

The potential for significant revenue loss when systems are down for an extended period of time is readily understood in today's business environment, because the majority of businesses rely exclusively on systems for much of their information needs. However, the cost of replacing systems and information in the event of catastrophic loss is often grossly underrated. Major organisations, when queried on insurance coverage for systems, come up with some surprising answers. Typically, organisations have coverage for mainframes and midrange systems and for the software for these environments. The workstations and the network servers, however, are often deemed as not valuable enough to insure. Coverage for the information itself is usually neglected as well, despite the fact that the major replacement cost for a company in crisis is the recreation of its information data base.

Notably, the personal computer, regardless of how it is configured or networked, is usually perceived as a standalone unit from the risk assessment point of view. Even companies that have retired their mainframes and embraced extensive client/server architecture, and that fully comprehend the impact of the loss of its use, erroneously consider only the replacement cost of the unit rather than that of the distributed system as the basis of risk.

Risk assessment is the control point of the recovery planning process. The amount of exposure a company believes it has, or is willing to accept, determines how much effort the company will expend on this process. Simply put, a company with no plan is fully exposed to catastrophic loss. Companies developing plans must approach risk assumption by identifying their worst-case scenario and then deciding how much they will spend to offset that scenario through mitigation, contingency plans, and training. Risk assessment is the phase required to formulate a company's management perspective, which in turn supports the goal of developing and maintaining a companywide contingency plan.

• Mitigation

The primary objectives of mitigation are to lessen risk exposures and to minimize possible losses. History provides several lessons in this area. For example, since the underground floods of 1992, companies in Chicago think twice before installing data centers in the basements of buildings. Bracing key computer equipment and office furniture has become popular in California because of potential injuries to personnel and the threat of loss of assets from earthquakes. Forward-thinking companies in the South and southern Atlantic states are installing systems far from the exterior of buildings because of the potential damage from hurricanes.

Although it is a simple exercise to make a backup copy of key data and systems, it is difficult to enforce this activity in a distributed systems environment. As systems have been distributed and the end user empowered, the regimen of daily or periodic backups has been adversely affected. In other words, the end user has been empowered with tools but has not been educated about, or held responsible for, the security measures that are required for those tools. One company, a leader in the optical disk-drive market, performs daily backups of its accounting and manufacturing systems to optical disk (using its own product), but never rotates the media and has never considered storing the backup off-site. Any event affecting the hardware (e.g., fire, theft, or earthquake) could therefore destroy the sole backup and the means of business recovery for this premier company. Mitigation efforts must counter such oversights.

• Preparation

The preparation phase of the disaster planning process delineates what specific actions must be taken should a disaster occur. Based on an understanding of plausible threats, planners must determine who will take what action if a disaster occurs. Alternates should be identified for key staff members who may have been injured as a result of the event. A location for temporary operations should be established in case the company's building is inaccessible after a disaster, and the equipment, supplies, and company records that will be required at this site should be identified. Preparation may include establishing a hot site for systems and telecommunications. Off-hours or emergency telephone numbers should be kept for all vendors and services providers that may need to be contacted. Moreover, the contingency plans must be clearly documented and communicated to all personnel.

• Testing

The testing phase proves the viability of the planning efforts. The recovery planner must determine, during testing, whether there are invalid assumptions and inadequate solutions in the company's plan. It is important to remember that organisations are not static and that an ever-changing business environment requires a reasonable frequency of testing. Recovery planners must repeat this phase of the plan until they are comfortable with the results and sure that the plan will work in a time of crisis.

• Response and Recovery

This final phase of the contingency plan is one that organisations hope never to have to employ. Preparing for actual response and recovery

includes identifying individuals and training them to take part in emergency response in terms of assessment of damage, cleanup, restoration, alternate site start-up, emergency operations duties, and any other activities that managing the crisis might demand.

Every phase of the planning process, prior to this phase, is based on normalcy. The planning effort is based on what is perceived to be plausible. Responses are developed to cover plausible crises and are done so under rational conditions. However, dealing with a catastrophic crisis is not a normal part of an employee's work day, and the recovery team must be tested under more realistic conditions to gauge how they will perform under stress and where lapses in response might occur. Ideally, recovery planners should stage tests that involve role playing to give their team members a sense of what they may be exposed to in a time of crisis.

3.2 Departmental Planning

Often, consultants are asked to help a company develop its business resumption plan and to focus only on the systems environment to reduce the overall cost of planning efforts. Often, companies take action on planning as the result of an information systems audit and thus focus solely on systems exposure and audit compliance. These companies erroneously view disaster recovery as an expense rather than as an investment in business continuity.

A plan that addresses data integrity and systems survivability is certainly a sound place to begin, but there are many other factors to consider in recovery planning. Depending on the nature of the business, for example, telecommunications availability may be much more important than systems availability. In a manufacturing environment, if the building and equipment are damaged in a disaster, getting the systems up and running may not necessarily be a top priority.

A company's business continuation plan should be a compilation of individual department plans. It is essential that each department identify its processes and prioritize those processes in terms of recovery. Companywide operating and recovery priorities can then be established by the company's management based on the input supplied by the departments. Information technology, as a service department to all other departments, will be better equipped to plan recovery capacity and required system availability based on this detailed knowledge of departmental recovery priorities.

Information Technology's Role

Information technology personnel should not be responsible for creating individual department plans, but they should take a leadership role in the plan development. Information technology generally has the best appreciation and understanding of information flow throughout the organisation. Its staff, therefore, are in the best position to identify and assess the following areas.

Interdepartmental Dependencies

It is common for conflicts in priorities to arise between a company's overall recovery plan and its departmental plans. This conflict occurs because departments tend to develop plans on their own without considering other departments. One department may downplay the generation of certain information because that information has little importance to its operations, but the same information might be vitally important to the operations of another department. Information technology departments can usually identify these discrepancies in priorities by carefully reviewing each department's plan.

External Dependencies

During the discovery process, recovery planners should determine with what outside services end-user departments are linked. End-user departments often think of external services as being outside the scope of their recovery planning efforts, despite the fact that dedicated or unique hardware and software are required to use the outside services. At a minimum, departmental plans must include the emergency contact numbers for these outside services and any company account codes that permit linkage to the service from a recovery location. Recovery planners should also assess the outside service providers' contingency plans for assisting the company in its recovery efforts.

Internal and External Exposures

Standalone systems acquired by departments for a special purpose are often not linked to a company's networks. Consequently, they are often overlooked in terms of data security practices. For example, a mortgage company funded all of its loans via wire transfer from one of three standalone systems. This service was one of the key operations of the company. Each system was equipped with a modem and a uniquely serialized encryption card for access to the wire service. However, these systems were not maintained by the information technology department, no data or system backups were maintained by the end-user department,

and each system was tied to a distinct phone line. Any mishap involving these three systems could have potentially put this department several days, if not weeks, in arrears in funding its loans. Under catastrophic conditions, a replacement encryption card and linkage establishment would have taken as much as a month to acquire.

As a result of this discovery, the company identified a secondary site and filed a standby encryption card, an associated alternate phone line, and a disaster recovery action plan with the wire service. This one discovery, and its resolution, more than justified the expense of the entire planning effort.

During the discovery process, the recovery planner identified another external exposure for the same company. This exposure related to power and the requirements of the company's uninterruptable power supply (UPS). The line of questioning dealt with the sufficiency of battery backup capacity and whether an external generator should be considered in case of a prolonged power interruption. An assumption had been made by the company that, in the event of an areawide disaster, power would be restored within 24 hours. The company had 8 hours of battery capacity that would suffice for its main operational shift. Although the county's power utility company had a policy of restoring power on a priority basis for the large employers of the county, the company was actually based in a special district and acquired its power from the city, not the county. Therefore, it would have power restored only after all the emergency services and city agencies were restored to full power. Moreover, no one could pinpoint how long this restoration period would be. To mitigate this exposure, the company added an external generator to its UPS system.

Appraise Management of Risks and Mitigation Costs

As an information technology department identifies various risks, it is the department's responsibility to make management aware of them. This responsibility covers all security issues—system survivability issues (i.e., disaster recovery), confidentiality, and system integrity issues.

In today's downsized environments, many information technology departments have to manage increasingly more complex systems with fewer personnel. Because of these organisational challenges, it is more important for the information technology staff involved in the planning process to present management with clear proposals for risk mitigation. Advocating comprehensive planning and security measures, and following through with management to see that they are implemented,

will ensure that a depleted information technology staff is not caught off-guard in the event of disaster.

Policies

To implement a system or data safeguard strategy, planners must first develop a policy—or standard operating procedure — that explains why the safeguard should be established and how it will be implemented. The planners should then get approval for this policy from management.

In the process of putting together a disaster recovery plan for a community college's central computing operations, one recovery planner discovered that numerous departments had isolated themselves from the networks supported by the information technology group. These departments believed that the servers were always crashing, which had been a cause for concern in years past, and they chose to separate themselves from the servers for what they considered to be safer conditions. These departments, which included accounting, processed everything locally on hard drives with no backups whatsoever. Needless to say, a fire or similar disaster in the accounting department would severely disrupt, if not suspend, the college's operations.

The recovery planner addressed this problem with a fundamental method of distributed system security: distribute the responsibility of data integrity along the channels of distributed system capability. A college policy statement on data integrity was developed and issued to this effect. The policy outlined end-user security responsibilities, as well as those of the department administrators.

Establish Recovery Capability

Based on departmental input and a company's established priorities, the information technology department must design an intermediate system configuration that is adequately sized to permit the company's recovery immediately following the disaster. Initially, this configuration, whether it is local, at an alternate company site, or at a hot site, must sustain the highest-priority applications yet be adaptable to addressing other priorities. These added needs will arise depending on how long it takes to reoccupy the company's facilities and fully restore all operations to normal. For example, planners must decide that the key client/server applications are critical to company operations, whereas office automation tools are not.

Restore Full Operational Access

The information technology department's plan should also address the move back from an alternate site and the resources that will be required to restore and resume full operations. Depending on the size of the enterprise and the plausible disaster, this could include a huge number of end-user workstations. At the very least, this step is as complex as a company's move to a new location.

3.3 Planning For the Distributed Environment

First and foremost, planners in a distributed environment must define the scope of their project. Determining the extent of recovery is the first step. For example, will the plan focus on just the servers or on the entire enterprise's systems and data? The scope of recovery, the departmental and company priorities, and recovery plan funding will delimit the planner's options. The following discussion outlines the basics of recovery planning regardless of budget considerations.

Protecting the LAN

Computer rooms are built to provide both special environmental conditions and security control. Environmental conditions include air conditioning, fire-rated walls, dry sprinkler systems, special fire abatement systems (e.g., Halon, FM-200), raised flooring, cable chase-ways, equipment racking, equipment bracing, power conditioning, and continuous power (UPS) systems. Control includes a variety of factors: access, external security, and internal security. All these aspects of protection are built-in benefits of the computer room. Today, however, company facilities are distributed and open; servers and network equipment can be found on desktops in open areas, on carts with wheels, and in communications closets that are unlocked or have no conditioned power. Just about anything and everything important to the company is on these servers or accessible through them.

Internal Environmental Factors: A computer room is a viable security option, though there are some subtleties to designing one specifically for a client/server environment. If the equipment is to be rack mounted, racking can be suspended from the ceiling, which yields clearance from the floor and avoids possible water damage. Notably, the cooling aspects of a raised floor design, plus its ability to hide a morass of cabling, are no longer needed in a distributed environment.

Conditioned power requirements have inadvertently modified computer room designs as well. If an existing computer room has a shunt trip by the exit but small standalone battery backup units are placed on servers, planners must review the computer room emergency shutdown procedures. The function of the shunt trip was originally to kill all

power in the room so that, if operational personnel had to leave in a hurry, they would be able to come back later and reset systems in a controlled sequence. Now, when there are individual battery backup units that sustain the equipment in the room, the equipment will continue to run after the shunt is thrown. Rewiring the room for all wall circuits to run off the master UPS, in proper sequence with the shunt trip, should resolve this conflict.

Room placement within the greater facility is also a consideration. When designing a room from scratch, planners should identify an area with structural integrity, avoid windows, and eliminate overhead plumbing.

Alternate fire suppression systems are still a viable protection strategy for expensive electronics and the operational on-site tape backups within a room. If these systems are beyond the company's budget, planners might consider multiple computer rooms (companies with a multiple-building campus environment or multiple locations can readily adapt these as a recovery strategy) with sprinklers and some tarpaulins handy to protect the equipment from incidental water damage (e.g., a broken sprinkler pipe). A data safe may also be a worthwhile investment for the backup media maintained on-site. However, if the company uses a safe, its personnel must be trained to keep it closed. In eight out of ten site visits where a data safe is used, the door is kept ajar (purely as a convenience). The safe only protects the company's media when it is sealed. If the standard practice is to keep it closed, personnel will not have to remember to shut it as they evacuate the computer room under the stress of an emergency.

If the company occupies several floors within a building and maintains communication equipment (e.g., servers, hubs, or modems) within closets, the closets should be treated as miniature computer rooms. The doors to the closets should be locked, and the closets should be equipped with power conditioning and adequate ventilation.

Physical Security: The other priority addressed by a properly secured computer room is control: control of access to the equipment, cabling, and backup media. Servers out in the open are prime targets for mishaps ranging from innocent tampering to outright theft. A thief who steals a server gets away not only with an expensive piece of equipment but with a wealth of information that may prove to be much more valuable and marketable than the equipment itself.

The college satellite campus, discussed earlier, had no backup of the information contained within its network. The recovery planner explained to the campus administration, which kept its servers out in the

open in its administration office area (a temporary trailer), that a simple theft of the $2,000 equipment would challenge its ability to continue operations. All student records, transcripts, course catalogs, instructor directories, and financial aid records were maintained on the servers. With no backup to rely on and its primary source of information evaporated, the campus administration would be faced with literally thousands of hours of effort to reconstruct its information base.

Property Management: Knowing what and where the organisation's computer assets (i.e., hardware, software, and information) are at any moment is critical to recovery efforts. The information technology department must be aware of not only the assets within the computer room but of every workstation used throughout the organisation: whether it is connected to a network (including portables), what its specific configuration is, what software resides on it, and what job function it supports. This knowledge is achievable if all hardware and software acquisitions and installations are run through the IT department, if the company's policies and procedures support information technology's control (i.e., all departments and all personnel willingly adhere to the policies and procedures), and if the department's property management inventory is properly maintained. Size is also a factor here. If the information technology department manages an organisation with a single server and 50 workstations, the task may not be too large; however, if it supports several servers and several hundred workstations, the amount of effort involved is considerable.

Data Integrity: Information, if lost or destroyed, is the one aspect of a company's systems that cannot be replaced simply by ordering another copy or another component. The company may have insurance, hot-site agreements, or quick-replacement arrangements for hardware and global license agreements for software, but its data integrity process is entirely in the hands of its information technology specialists. The information technology specialist and the disaster recovery planner are the individuals who must ensure that the company's information will be recoverable.

Based on the initial risk assessment phase, planners can determine just how extensive the data integrity program should be. The program should include appropriate policies and education addressing frequency of backups, storage locations, retention schedules, and the periodic verification that the backups are being done correctly. If the planning process has just begun, data integrity should be the first area on which planners focus their attention. None of the other strategies they implement will count if no means of recovering vital data exist.

Network Recovery Strategies

The information technology specialist's prime objective with respect to systems contingency planning is system survivability. In other words, provisions must be in place, albeit in a limited capacity, that will support the company's system needs for priority processing through the first few hours immediately following a disaster.

Fault Tolerance vs. Redundancy: To a degree, information technology specialists are striving for what is called fault tolerance of the company's critical systems. Fault tolerance means that no single point of failure will stop the system. Fault tolerance is often built in as part of the operational component design of a system. Redundancy, or duplication of key components, is the basis of fault tolerance. When fault tolerance cannot be built in, a quick replacement or repair program should be devised. Moving to an alternate site (i.e., a hot site) is one quick replacement strategy.

Alternate Sites and System Sizing: Once the recovery planner fully understands the company's priorities, the planner can size the amount of system capacity required to support those priorities in the first few hours, days, and weeks following a disaster. When planning for a recovery site or establishing a contract with a hot-site service provider, the information technology specialist must size the immediate recovery capacity. This is extremely important, because most hot-site service providers will not allow a company to modify its requirements once it has declared a disaster.

The good news with respect to distributed systems is that hot-site providers offer options for recovery. These options often include offering the use of their recovery center, bringing self-contained vans to the company's facility (equipped with the company's own secured configuration), or shipping replacement equipment for anything that has been lost.

Adequate Backups with Secure Off-Site Storage: This process must be based on established company policies that identify vital information and detail how its integrity will be managed. The work flow of the company and the volatility of its information base dictate the frequency of backups. At a minimum, backup should occur daily for servers and weekly or monthly for key files of individual workstations.

Planners must decide when and how often to take backups for off-site. Depending on a company's budget, off-site could be the building next door, a bank safety deposit box, the network administrator's house, the branch office across town, or a secure media vault at a storage facility maintained by an off-site media storage company. Once the company

meets the objective of separating the backup copy of vital data from its source, it must address the accessibility of the off-site copy.

The security of the company's information is of vital concern. The planner must know where the information are to be kept and about possible exposure risks during transit. Some off-site storage companies intentionally use unmarked, nondescript vehicles to transport a company's backup tapes to and from storage. These companies know that this information is valuable and that its transport and storage place should not be advertised.

Adequate LAN Administration: Keeping track of everything the company owns - its hardware, software, and information bases - is fundamental to a company's recovery effort. The best aid in this area is a solid audit application that is run periodically on all workstations. This procedure assists the information technology specialist in maintaining an accurate inventory across the enterprise and provides a tool for monitoring software acquisitions and hardware configuration modifications. The inventory is extremely beneficial for insurance loss purposes. It also provides the technology specialist with accurate records for license compliance and application revision maintenance.

Personnel: Systems personnel are too often overlooked in systems recovery planning. Are there adequate systems personnel to handle the complexities of response and recovery? What if a key individual is affected by the same catastrophic event that destroys the systems? This event could cause a single point of failure.

An option available to the planner is to propose an emergency outsourcing contract. A qualified systems engineer hired to assist on a key project that never seems to get completed (e.g., the network system documentation) may be a cost-effective security measure. Once that project is completed to satisfaction, the company can consider structuring a contractual arrangement that, for example, retains the engineer for one to three days a month to continue to work on documentation and other special projects, as well as cover for staff vacations and sick days, and guarantees that the engineer will be available on an as-needed basis should the company experience an emergency. The advantage of this concept is that the company maintains effective outsourced personnel who are well versed in the company's systems if the company needs to rely on them during an emergency.

3.4 Testing

The success of a business recovery plan depends on testing its assumptions and solutions. Testing and training keep the plan up-to-date and maintain the viability of full recovery.

Tests can be conducted in a variety of ways: from reading through the plan and thinking through the outcome to full parallel system testing, or setting up operations at a hot site or alternate location and having the users run operations remotely. The full parallel system test generally verifies that the hot-site equipment and remote linkages work, but it does not necessarily test the feasibility of the user departments' plans. Full parallel testing is also generally staged within a limited amount of time, which trains staff to get things done correctly under time constraints.

Advantages of the Distributed Environment for Testing

Because of their size and modularity, distributed client/server systems provide a readily available, modifiable, and affordable system setup for testing. They allow for a testing concept called cycle testing.

Cycle testing is similar to cycle counting, a process used in manufacturing whereby inventory is categorized by value and counted several times a year rather than in a one-time physical inventory. With cycle counting, inventory is counted year long, with portions of the inventory being selected to be counted either on a random basis or on a preselected basis. Inventory is further classified into categories so that the more expensive or critical inventory items are counted more frequently and the less expensive items less frequently. The end result is the same as taking a one-time physical inventory in that, by the end of a calendar year, all the inventory has been counted. The cycle counting method has several advantages:

- Operations do not have to be completely shut down while the inventory is being taken.
- Counts are not taken under time pressure, which results in more accurate counts.
- Errors in inventories are discovered and corrected as part of the continuous process.

The advantages of cycle testing are similar to those of cycle counting. Response and recovery plan tests can be staged with small manageable groups so they are not disruptive to company operations. Tests can be staged by a small team of facilitators and observers on a continual basis. Tests can be staged and debriefings held without time pressure, allowing the participants the time to understand their roles and the planners the time to evaluate team response to the test scenarios and to make

necessary corrections to the plan. Any inconsistencies or omissions in a department's plan can be discovered and resolved immediately among the working participants.

Just as more critical inventory items can be accounted for on a more frequent basis, so can the crucial components required for business recovery (i.e., systems and telecommunications). With the widespread use of LANs and client/server systems, information systems departments have the opportunity to work with other departments in testing their plans.

4.0 CONCLUSION

Developing a business recovery plan is not a one-time, static task. It is a process that requires the commitment and cooperation of the entire company. To perpetuate the process, business recovery planning must be a company-stipulated policy in addition to being a company-sponsored goal. Organisations must actively maintain and test plans, training their employees to respond in a crisis. The primary objective in developing a business resumption plan is to preserve the survivability of the business.

An organisation's business resumption plan is an orchestrated collection of departmental responses and recovery plans. The information technology department is typically in the best position to facilitate other departments' plan developments and can be particularly helpful in identifying the organisation's interdepartmental information dependencies and external dependencies for information access and exchange.

A few protective security measures should be fundamental to the information technology department's plan, no matter what the scope of plausible disasters. From operational mishaps to area wide disasters, recovery planners should ensure that the information technology department's plan addresses:

- An adequate backup methodology with off-site storage.
- Sufficient physical security mechanisms for the servers and key network components.
- Sufficient logical security measures for the organisation's information assets.
- Adequate LAN/WAN administration, including up-to-date inventories of equipment and software.

Finally, in support of an organisation's goal to have its business resumption planning process in place to facilitate a quick response to a

crisis, the plan must be sufficiently and repeatedly tested, and the key team members sufficiently trained. When testing is routine, it becomes the feedback step that keeps the plan current, the response and recovery strategies properly aligned, and the responsible team members ready to respond. Testing is the key to plan viability and thus to the ultimate survival of the business.

5.0 SUMMARY

- Today's organisations, in their efforts to reduce costs, are streamlining layers of management while implementing more complex matrices of control and reporting.
- Distinctive areas must be addressed in the formulation of a company's disaster recovery plan, and attention to these areas should follow the steps of the scientific method: a statement of the problem, the development of a hypothesis, and the testing of the hypothesis.
- Often, consultants are asked to help a company develop its business resumption plan and to focus only on the systems environment to reduce the overall cost of planning efforts.
- To implement a system or data safeguard strategy, planners must first develop a policy—or standard operating procedure—that explains why the safeguard should be established and how it will be implemented.
- First and foremost, planners in a distributed environment must define the scope of their project. Determining the extent of recovery is the first step.
- The success of a business recovery plan depends on testing its assumptions and solutions. Testing and training keep the plan up-to-date and maintain the viability of full recovery.
- Cycle testing is similar to cycle counting, a process used in manufacturing whereby inventory is categorized by value and counted several times a year rather than in a one-time physical inventory

6.0 TUTOR-MARKED ASSIGNMENT

1. List the phases of a disaster recovery plan process.
2. Discuss briefly the advantages of the Distributed Environment for Testing.

7.0 REFERENCES/FURTHER READINGS

Krause, M. and Tipton, H.F. Handbook of Information Security *Management.*

UNIT 2 INFORMATION SYSTEMS SECURITY CONTROLS

CONTENTS

1.0 INTRODUCTION

Systems controls are defense mechanisms intended to prevent accidental hazards, deter intentional acts, detect problems as early as possible, enhance damage recovery, and correct problems. Security controls are designed to protect all the components of an information system, specifically data, software, hardware, and networks.

Knowing about major potential threats to information systems of an e-business is important, but understanding ways to defend against these threats is equally critical. Defending information systems is not a simple or inexpensive task for the following reasons:

- Hundreds of potential threats exist
- Computing resources may be situated in many locations
- Many individuals control information assets
- Computer networks can be outside the organisation and difficult to protect
- Rapid technological changes make some controls obsolete as soon as they are installed
- Many computer crimes are undetected for a long period of time, so it is difficult to “learn from experience”.
- People tend to violate security procedures because the procedures are inconvenient

- Many computer criminals who are caught goes unpunished, so there is no deterrent effect
- The amount of computer knowledge necessary to commit computer crimes is usually minimal. As a matter of fact one can learn hacking for free on the Internet
- The costs of preventing hazards can be very high. Therefore most organisations simply cannot afford to protect against all possible hazards
- It is difficult to conduct cost-benefit justification for controls before an attack occurs since it is difficult to assess the value of a hypothetical attack.

Therefore organizing an appropriate defense system is one of the major activities of any prudent information system or functional manager who controls information resources.

Protecting information systems in e-businesses is accomplished by inserting controls. Controls can be integrated into hardware and software during the systems development phase (a most efficient approach). They can also be implemented once the system is in operation or during its maintenance. The important point is that defense should stress prevention, it does no good after the crime. Since there are many threats there are also many defense mechanisms.

2.0 OBJECTIVES

At the end of this unit you should be able to:

- explain why defending information systems are not easy
- identify the major defense strategies in dealing with control of systems
- explain the different types of controls and to differentiate them
- understand how to implement controls generally
- understand how to put in place a disaster recovery plan in the cases of emergencies.

3.0 MAIN CONTENT

3.1 Defense Strategies: How to Protect

The selection of a specific strategy depends on the objective of the defense and on the perceived cost-benefit. The following are the major defense strategies:

1) **Controls for Prevention and Deterrence. Properly designed** controls may prevent errors from occurring, deter criminals from

attacking the system, and better yet, deny access to the unauthorized people. Prevention and deterrence are especially important where the potential damage is very high.

2) Detection. It may not be economically feasible to prevent all hazards, and deterring measures might not work. Therefore, unprotected systems are vulnerable to attack. Like fire, the earlier it is detected, the easier it is to combat and the less is the damage. Detection can be performed in many cases by using special diagnostic software.

3) Limitation. This means to minimize losses once a malfunction has occurred. Users typically want their systems back in operation as quickly as possible. This can be accomplished by including a fault-tolerant system that permits operation in a degraded mode until full recovery is made. If a fault-tolerant system does not exist, a quick (and possibly expensive) recovery must take place.

4) Recovery. A recovery plan explains how to fix a damaged information system as quickly as possible. Replacing rather than repairing components is one route to fast recovery.

5) Correction. Correcting damaged systems can prevent the problem from occurring again.

The defense strategy may involve the use of several controls as described in subsequent parts of this unit.

Information systems controls can be divided into two major groups: general (system) control and application controls.

3.2 General Controls

General controls are established to protect the system regardless of the specific applications. For example, protecting hardware and controlling access to the data centre are independent of the specific application.

The major categories of general controls are physical controls, data security controls, communications (network) controls and administrative controls.

Physical Controls

Physical security refers to the protection of computer facilities and resources. This includes protecting physical property such as computers, data centres, software manuals, and networks. Physical security is the first line of defense and usually the easiest to construct. It provides

protection against most natural hazards as well as against some human hazards. Appropriate physical security may include several controls such as the following:

- Appropriate design of the data centre. For example, the site should be non-combustible and waterproof.
- Shielding against electromagnetic fields.
- Good fire prevention, detection, and extinguishing systems, including sprinkler system, water pumps, and adequate drainage facilities. A better system is a fire-enveloping Halon gas system.
- Emergency power shutoff. And backup batteries that must be maintained in operational condition.
- Properly designed, maintained, and operated air-conditioning systems
- Motion detector alarms that deter physical intrusion.

Access Control

Access control is the restriction of unauthorized user access to a portion of a computer system or to the entire system. To gain access, a user must first be authorized. Then when the user attempts to gain access, he or she must be authorized.

Access to a computer system basically consist of three steps:

1. Physical access to a terminal
2. Access to the system, and
3. Access to specific commands, transactions, privileges, programs and data within the system.

Access control software is commercially available for large mainframes, microcomputers, personal computers, local area networks, and dial-in communications network. Access control to the network is executed through firewalls.

Access procedures match every valid user with a unique user identifier (UID). They also provide an authorization method to verify who they claim to be. User identification can be accomplished when the following identifies each user:

- Something only the user knows, such as password
- Something only the user has, for example smart card or a token
- Something only the user is, such as a signature, voice, fingerprints, or retinal (eye) scan. It is implemented through biometric controls.

Biometric control is defined as an "automated method of verifying the identity of a person, based on psychological or behavioral characteristics". The most common biometric are the following:

- Photo. The computer takes a picture of your face and matches it with a prestored picture. In 1997, this method was successful in correctly identifying users except in cases of identical twins.
- Fingerprints. Each time a user wants access, matching a fingerprint against a template containing the authorized person's fingerprint identifies him or her.
- Hand geometry. Similar to fingerprints except the verifier uses a television-like camera to take a picture of the user's head. Certain characteristics of the hand (e.g. finger length and thickness) are electronically compared against the information stored in the computer.
- Blood vessel pattern in the retina of a person's eye. A match is attempted between the pattern of the blood vessels ion the back-of-the-eye retina that is being scanned and a prestored picture of the retina.
- Voice. A match is attempted between the user's voice and the voice pattern stored on templates.
- Signature. Signatures are matched against the prestored authentic signature. This method can supplement a photo-card ID system.
- Keystroke dynamics. A match of the person's keyboard pressure and speed against prestored information.
- Others. Several other methods exist such as facial thermography and iris scan.

Data Security Control

Data security is concerned with protecting data from accidental or intentional disclosure to unauthorized persons, or from unauthorized modification or destruction. Data security functions are implemented through operating systems, security access programs, database/data communications products, recommended backup/recovery procedures, application programs, and external control procedures. Data security must address the following issues: confidentiality of data, access *control, critical nature of data, and integrity of data.*

Two basic principles should be reflected in data security.

- Minimal Privilege. Only the information a user needs to carry out an assigned task should be made available to him or her.
- Minimal Exposure. Once a user gains access to sensitive information, he or she has the responsibility of protecting it by

making sure only people whose duties require it obtain knowledge of the information while it is processed, stored or in transit.

Data integrity is the condition that exists as long as accidental or intentional destruction, alteration, or loss of data does not exist or occur. It is the preservation of data for its intended use.

Communication (Network) Control

Network protection is becoming extremely important as the use of the Internet, intranets, and electronic business /commerce increases. This is discussed in details in other units of this course.

Administrative Controls

While the previously discussed controls were technical in nature, administrative controls deal with issuing guidelines and monitoring compliance with the guidelines. Representative examples of such controls include the following:

- Appropriately selecting, training and supervising employees, especially in accounting and information systems.
- Fostering company loyalty
- Immediately invoking access privileges of dismissed, resigned or transferred employee
- Requiring periodic modification of access controls (such as password)
- Developing programming and documentation standards (to make auditing easier and to use the standard as guide for employees).
- Insisting on security bonds or malfeasance insurance for key employees
- Instituting separation of duties, namely dividing sensitive computer duties among as many employees as economically feasible in order to decrease the chance of intentional or unintentional damage.
- Holding periodic random audits of the system.

Other General Controls

Several other types of controls are considered general. Representative examples include the following:

- Programming Controls. Errors in programming may result in costly problems. Causes include the use of incorrect algorithms or programming instructions, carelessness, inadequate testing and configuration management, or lax security. Controls include training, establishing standards for testing and configuration management, and enforcing documentation standards.

• Misunderstanding or Misinterpretations. Manuals are often a source of problems because they are difficult to interpret or may be out of date. Accurate writing, standardization updating, and testing are examples of appropriate documentation control. Intelligent agents can be used to prevent such problem.

• Systems Development Controls. Systems development controls ensure that a system is developed according to established policies and procedures. Conformity with budget, timing, security measures, and quality and documentation requirements must be maintained.

3.3 Application Controls

Application controls are safeguards that are intended to protect specific applications.

General controls are intended to protect the computing facilities and provide security for hardware, software, data, and networks. However, general controls do not protect the content of each specific application. Therefore, controls are frequently built into the applications (that is, they are part of the software) and are usually written as validation rules. They can be classified into three major categories: input controls, *programming controls and output controls.*

Input Controls

Input controls are designed to prevent data alteration or loss. Data are checked for accuracy, completeness, and consistency. Input controls are very important; they prevent the GIGO (garbage-in-garbage-out) situation. Examples of input controls are the following:

• Completeness. Items should be of a specific length (e.g. nine digits for a Social Security number). Addresses should include a street, city, state and zip code.
• Format. Formats should be a standard form. For example, sequences must be preserved (zip code comes after an address).
• Range. Only data within a specified range are acceptable. For example, zip code ranges between 10,000 to 99,999, the age of a person cannot be larger than say, 120 and hourly wage cannot exceed %50.
• Consistency. Data collected from two or more sources need to be matched. For example, in medical history data males cannot be pregnant.

Processing Controls

Processing controls ensure that data are complete, valid, and accurate when being processed and that program have been properly executed. These programs allow only authorized users to access certain programs or facilities and monitor the complete use by individuals.

Output Controls

Output controls ensure that the results of computer processing are accurate, valid, complete, and consistent. By studying the nature of common output errors, management can evaluate possible controls to deal with problems. Also, controls ensure that outputs are sent only to authorize personnel.

3.4 Implementing Controls

Implementing controls in an organisation can be very complicated tasks, particularly in large, decentralized companies where administrative controls may be difficult to enforce. For example, Lee (1990) suggests that IT auditing be expanded to include end-user computing. Of the many issues involved in implementing controls, two are described here: planning and organizing, and disaster recovery planning.

Planning and Organizing

A comprehensive control and security management program must begin with the establishment of a formal documented organisation wide secure policy endorsed by the highest levels of management in the organisation. Such a program was developed by Fine (1983) Fine's total computer security can be envisioned as a horizontal beam supported by nine pillars. Three of the pillars are technical issues: physical security, *control and system security. These three support the integrity and* confidentiality of an organisation's information systems.

The other six pillars are:

- A defined and documented computer security policy
- Standards and procedures
- An assignment of responsibility for computer security
- A personal security program
- A complete asset-threat inventory
- Introduction of user awareness.

Each of the program's pillars must be carefully planned and properly managed. This may not be a simple task, since the pillars are heavily

Developing a security policy begins with a detailed analysis of current equipment, functions performed, data contained, ease of access, security devices and potential losses. Following this analysis, a policy can be drafted that will focus on employees as well as other assets (hardware, software and data). Any security policy developed should deal with threats proactively, not just reactively. For each identified threat, the policy should describe not only the necessary protection, but also the action to be taken if the threat materializes. Stringent policies, surprise audits, and strenuous security awareness programs are excellent ways of protecting the company information systems.

Security administration is the responsibility of both the information systems department and the line managers. Line managers are responsible for protecting all resources in their possession and for ensuring that their subordinates are aware of and abide by established security policies and procedures.

3.6 Disaster Recovery Planning

Disaster may occur in many places without warning. According to Strassman (1997), the best defense is to be prepared. Therefore, an important element in any security system is the disaster recovery plan. Destruction of all (or some) of the computing facilities in an e-business for example, can cause significant damage. Therefore it is difficult for many organisations to obtain insurance for their computers and information system without showing satisfactory disaster prevention and recovery plan.

Disaster recovery is the chain of events linking planning to protection to recovery. The following are the key thoughts by Knoll (1986):

- The purpose of recovery planning is to keep the business running after a disaster occurs. Both the information systems department and the line management should be involved in the preparation of the plan. Each function in the business should have a valid recovery capability plan.
- Recovery planning is part of asset protection. Every organisation should assign responsibility to management to identify and protect assts within their sphere of functional control.
- Planning should focus first on recovery from a total loss of all capabilities
- Proof of capability often involves some kind of what-if analysis that shows that the recovery plan is current.
- The plan should be kept in a safe place; copies should be given to key managers and the plan should be audited periodically.

- All critical applications must be identified and their recovery procedures addressed in the plan.
- The plan should be written so that it should be effective incase of disaster, not just in order to satisfy the auditors.

Disaster recovery planning can be complex, and it may take several months to complete (Butler, 1992). Using special software, the planning job can be expedited. See Study Case 1.

Study Case 1: PCBASED SOFTWARE PROVIDES USEFUL DISASTER RECOVERY

Hurricane Hugo hit Charleston, South Carolina on Friday, September 20, 1989. Electrical power was knocked out for nine days. However, Heritage Trust Credit Federal (a credit union) resumed normal operation by the following Tuesday.

The credit union had developed a disaster recovery plan supported by battery-operated, PC-based software (Recovery Pac). Both internal and external auditors were involved in the plan. When a disaster occurred, the software was to be activated. The software tracks the crisis and suggests the best method of handling it. The software is flexible enough to guide users through problems that have not been anticipated in the test or the original plan.

Disaster Avoidance

Disaster avoidance is an approach oriented toward prevention. The idea is to minimize the chance of avoidable disasters (such as fire, or other human threats) for example, many organisations use a device called uninterrupted power supply (UPS) which provides power in case of power outage.

Backup Arrangements

In the event of a major disaster, it is often necessary to move a centralized computing facility to another backup location. External hot-site *vendors provide access to a fully configured backup data centre. To* appreciate the usefulness of such arrangement, consider the following examples:

On the evening of October 17, 1889, a major earthquake hit San Francisco, California, and Charles Schwab and Company was ready. Within a few minutes, the company's disaster plan was activated. Programmers, engineers, and backup computer tapes of October 17th

transactions were flown on a chartered jet to Carlstadt, New Jersey. There, Comdisco Disaster Recovery Service provided a hot-site. The next morning, the company resumed normal operation. Montgomery Securities, on the other hand, had no backup recovery plan. On October 18, the day after the quake, the traders had to use telephone rather than computer to execute trades. Montgomery lost revenues of $250,000 to $500,000 in one day.

A less costly alternative arrangements external cold-site vendors that provide empty office space with special flooring, ventilation, and wiring. In emergency, the stricken company moves its own (or leased) computers to the site.

Physical computer security is an integral part of a total security system. Cray Research a leading manufacturer of supercomputers has incorporated a corporate security plan, under which the corporate computers are monitored automatically and controlled centrally. Graphic displays show both normal status and disturbances. All the devices controlled are represented as icons on floor-plan graphics. These icons can change colours (green means normal, red signifies a problem). The icon can flash as well. Corrective action massages are displayed whenever appropriate. The alarm system includes over 1,000 alarms. Operators can be alerted, even at remote locations, in less than one second.

***Study Case 2: DISASTER PLANNING AND THE INTERNET AT* REUTERS LTD**

Reuters is a multinational information delivery corporation with great interest in the Web. One of its subsidiaries, Realty Online builds and operates online stock trading services fro well-known brokerage houses, such as Paine Webber, Inc. If Reuters's information system were to fail outright, it would take more than 15 brokerage houses with it. The cost, not to mention the legal ramifications, can be tremendous. There is a very little experience in how to plan for Web-based disaster (as at the time of this case). Both traditional vendors, such as Comdisco Inc and Internet-based specialists, such as Exodus Communications, are trying to grab a share of this developing market. Reuter implemented an Internet disaster recovery plan with SunGard Corp. in addition, the company operates three redundant web sites in different locations from coast to coast. In case all of them were to fail, a hot sit, which is constantly updated with information from the main Reuters server, would be used to assist continuous operation.

Source: Condensed from Interactive Work, March, 1998

4.0 CONCLUSION

A careful balance between preventive and detective control measures is needed to ensure that users consider the security controls reasonable and

to ensure that the controls do not overly inhibit productivity. The combination of physical, technical, and administrative controls best suited for a specific computing environment can be identified by completing a quantitative risk analysis. Because this is usually an expensive, tedious, and subjective process, however, an alternative approach—referred to as meeting the standard of due care—is often used. Controls that meet a standard of due care are those that would be considered prudent by most organisations in similar circumstances or environments.

5.0 SUMMARY

- Systems controls are defense mechanisms intended to prevent accidental hazards, deter intentional acts, detect problems as early as possible, enhance damage recovery, and correct problems.
- The selection of a specific strategy depends on the objective of the defense and on the perceived cost-benefit
- General controls are established to protect the system regardless of the specific applications. For example, protecting hardware and controlling access to the data centre are independent of the specific application.
- Application controls are safeguards that are intended to protect specific applications
- General controls are intended to protect the computing facilities and provide security for hardware, software, data, and networks
- Implementing controls in an organisation can be very complicated tasks, particularly in large, decentralized companies where administrative controls may be difficult to enforce.
- Disaster may occur in many places without warning. According to Strassman (1997), the best defense is to be prepared. Therefore, an important element in any security system is the disaster recovery plan.

6.0 TUTOR-MARKED ASSIGNMENT

List 10 reasons why it is difficult to defend information systems.

7.0 REFERENCES/FURTHER READINGS

Turban, E., McLean, E., and Wetherbe, J. (1999). Information *Technology for Management. John Wiley & Sons.*

UNIT 3 INFORMATION SYSTEMS AUDIT

CONTENTS

1.0 INTRODUCTION

Public auditors obtain part of the needed data from computerized accounting and Management systems. Auditors need to evaluate audit risk and reliability of the received information (audit evidence), therefore it is important to know how the audited entity controls information systems. Information received by auditors is not primary; it is obtained after a complex process of data processing during which errors may occur.

Errors may be made due to human factor, e.g., when entering data, due to programmers' errors etc. Errors may be random and intentional. Like any other assets, IS are vulnerable, e.g., they may be damaged or stolen. Data and programmes which are in the computer are intangible, therefore they may be accessed or changed without leaving any visible trace.

IS development, installation, and maintenance costs should also be properly audited. It is purposeful to evaluate economy, efficiency, and effectiveness of IS development, installation, and maintenance. Also, audit methods have to be continually developed taking into consideration progress of science and technology.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define information systems audit
- identify the different types of information systems audit and how they differ from each other
- answer the question of "What are processes of evaluation of internal Control information systems?"
- identify processes in the Performance of application controls audit
- explain how to carry out Information System Audit Planning.

3.0 MAIN CONTENT

3.1 General Principles

3.1.1 Definition of IS Audit

IS audit is a process of evidence collection and evaluation allowing whether a computer system (information system) ensures assets'security, data integrity, as well as helps to efficiently seek organisational goals and rationally use the resources.

3.1.2 Areas of IS Audit

Application of IS audit may be divided into two areas:

- Evaluation of internal control,
- Evaluation of IS in terms of economy, efficiency, and effectiveness (hereinafter–3Es). Auditing Requirements, Point 70.

Internal control shall mean the entire set of controls established by the management of a public legal entity in order to provide reasonable assurance that the operations of the public legal entity are legal, economic, efficient, effective and transparent, that the strategic and other plans are implemented, that assets are safeguarded, that financial information and reporting are reliable and exhaustive, that contractual liabilities to third persons are satisfied and that all identified risks are managed. In order to avoid problems of IS management and security protection general IS control methods were developed. Generally IS audit is meant to evaluate such control.

Evaluation of the audited entity's internal control is an area of financial and performance audits; therefore IS audit is a constituent part of financial and performance audits. Evaluation of IS in terms of economy,

efficiency, and effectiveness is a separate IS performance audit conducted following Performance Audit Manual.

3.1.3 Types and Objectives of IS Audit

Types of IS audit:

- Audit of IS general controls
- Audit of application controls
- Audit of IS development controls
- IS performance audit.

3.1.4 Objectives of IS Audit

- Audit of IS general controls is to evaluate internal control which covers all information systems of an organisation.
- Audit of application controls is to evaluate a control related to data input, processing, protection, and obtaining in the specific applications (e.g., Navision Financials, LABBIS etc.).
- Audit of IS development controls is to evaluate management and control of IS development from the beginning of its conception until its legitimation; covering IS change management;
- Objective of IS performance audit is to evaluate issues related to IS in terms of efficiency, economy, and effectiveness.

3.1.5 Cooperation of Public Auditors

In terms of IS audit INTOSAI distinguishes three levels of:

- Public auditors conducting financial and performance audits (hereinafter –general
- auditors),
- IS auditors,
- IS/IT specialists.

Generalist Auditor

Financial auditors use in their work computerised audit tools, carry out risk evaluation in non-complicated IS, having encountered problem relating to clients IS, consult IS auditors for more detailed analysis.
Performance auditors use in their work computerised audit tools, participate in analysing correctness, reliability and comprehensiveness of management information and evaluate IS from the point of view of economy, effectiveness and efficiency

IT/IS Specialist and IS Auditor

- Consultations of Specialists of IT Department

- Consultations of External IS/IT Specialists
- Helps to transfer data from the IS of audited entity to computerized audit tools
- Evaluates general controls of information system
- Helps to perform audits of application software for financial and performance auditors
- Participate in evaluating of IS from the point of view of economy, effectiveness and efficiency
- Prepares IS methodic and provides training
- Provides information to IT Department about good IT governance / management practice
- IS audits performed by general auditors are limited to medium complexity evaluation of IS general control and accounting programmes (e.g., Navision Financials, LABBIS etc.).
- IS auditors perform audits of general control of complex IS (e.g., IS of the State Social Insurance Fund Board of the Republic of Lithuania, Customs'IS etc.), IS development audits, and IS performance audits.
- IS/IT specialists provide specialized guidance on particular issues.

3.1.6 Types of IS Audit

During financial or performance audits, general auditors may ask for help from IS auditors or IS/IT specialists. In such cases Audit Department Director (Deputy Director) applies to Head of the structural unit of the NAOL which performs IS audits. In case of need external IT/ IS specialist may be invited. Having performed IS audits general auditors present results (copy of a document) of evaluation of IS general control to structural unit of the NAOL which performs IS audits.

IS audit is performed following:

- Public Auditing Requirements;
- Financial and Performance Audit Manuals;
- International standards of information systems audit and Control Association ISACA and guidelines of information system audit of this Association;
- European Implementing Guidelines for the INTOSAI Auditing Standards.

3.2 Process of Evaluation of IS Internal Control

Audit is performed upon an assignment following the procedure established by the Auditor General or as a separate stage of financial or

performance audit. Internal control evaluation of IS is performed in the following order:

- First of all, evaluation of IS general control is conducted in a particular audited entity, and maturity of IS general control is identified;
- Having performed evaluation of IS general control, evaluation of particular application software control is conducted;
- Annual summary on IS internal control is prepared;
- IS performance audit is carried out in the audited entity if potential IS economy, efficiency, and effectiveness problems are identified. Strategic planning of IS performance audit uses data obtained evaluating IS general control.

3.2.1 Audit of IS General Controls

Auditor should not consider computer processed and (or) transferred information of the audited entity reliable until he has proper supporting evidence. Such evidence could be obtained after getting assurance that internal control procedures of the system operate securely and properly.

Audit Resources of IS General Controls
Detailed evaluation of general control procedures for IS may require profound knowledge of IS audit and rather a lot of resources. Auditors who want to conduct a comprehensive audit of IS general control may need a specialist help; however, comprehensive IS audit is not always justifiable. General auditors need to decide when it is necessary to ask for help from IS audit specialists. To this end they need to use IS complexity evaluation questionnaire in which complexity evaluation criteria of information systems are pointed out. According to their complexity, information systems are divided into simple, medium complexity, and complex.

In audited entities which have introduced simple (not complex) IS, audit of the IS general control should be conducted by general auditors. Auditors should at once, without IS internal control evaluation, understand that simple (primitive) computerized accounting systems are subject to great risk.

Complex general internal control audit of the state's Information System is conducted by IS auditors.

Complex audits of the state's IS are included into the annual Public Audit Programme.

Performing of Audit

Audit of IS general control may be divided into three stages:

1. Analysis of IS general controls;
2. Testing of IS general controls;
3. Evaluation of IS internal controls.

Analysis of IS General Controls
In Evaluating IS general controls of an organisation an auditor should identify:

• Who in the management of an audited entity is responsible for IS;
Communicating with managers of various levels of an organisation it is purposefully to find out as to how the managers supervise IS processes, i.e., it is purposeful to get to know the monitoring system of an organisation and its operation. It is recommended to point out what actions were taken by managers when they had identified that IS performance had not satisfied legal acts and internal procedures. It is also purposeful to take note of the fact whether the applied control procedures are regularly (e.g., at least once a year) revised, and efficiency of the existing procedures as well as the need for new ones is regularly evaluated.

• How IT processes are organized in an audited entity;

• How internal audit of information systems is carried out (an auditor may use results of internal auditors if the quality of their work is reliable);

• Whether IS risk is evaluated in an audited entity; if there is an evaluation

• methodology, and whether this evaluation is documented;

• Whether IS and information security strategies, policies (regulations), procedures, rules, project specifications are documented, and whether an audited entity follows the above mentioned documents;

Auditors should learn which standards, methodologies, or other documents regulating to IT area are used by the audited entity. Legal acts of the Republic of Lithuania recommend following LST ISO/IEC 17799:2002 standard information technology. Practice code for information security management. (Equivalent to ISO/IEC 17799:2000)
"Audited entity may also choose LST ISO/IEC TR 13335 information technology. Guidelines for security management of standard

information technology (equivalent to ISO/IEC TR 13335:1996) "Standard questionnaires were developed for auditors according to the above mentioned standards. These questionnaires list standard questions (which may be freely chosen by an auditor). However, assessment of compliance with the above mentioned standards is rather complicated, therefore it is recommended to use help of IS auditors.

Identify whether IS comply with legal acts; Auditor has to analyse or identify the main indicators of IS activity (indicators, if the management of the audited entity has not identified such). Part of IS activity is regulated by legal acts, therefore is it recommended to use the questionnaire prepared according to legal acts of the Republic of Lithuania. This questionnaire presents standard questions. Legal acts and questions have to be chosen by an auditor taking into consideration the topic of public audit.

Evaluate other material information related to the use of information systems.

Auditor establishes that there is an inappropriate password policy in the organisation, i.e., employees has his own passwords but they are openly written on stickers which are placed on computers; colleagues use each others' passwords. Division of responsibilities is formal, in reality it is not implemented (an accountant uses his own password, senior accountant who controls the others has his password, however, it is known to his colleagues).

Auditor identifies that all the information (data bases) is contained in servers, however, backup copies are not made. Servers are placed in premises which are located under the sanitary unit, and usually there is a concrete ceiling.

• Testing of IS General Controls

Audit of IS general controls is not limited to the review of documents regulating assurance of internal control. Auditor has to make sure that control measures identified in legal acts, IS policy, procedures, rules, and any other documents are really operating.

To this end environment observation may be performed (e.g., observation of passing through the control post, entering the computer premises, workstations etc.) or interviews may be carried out (e.g., with managers of units, users who have different permissions, system administrator etc.). It is purposeful to make sure that having made

themselves familiar with control procedures employees understand and implement them, and provide reasoned suggestions on their development to the management. It is recommended to define potential situations emerging of which may be determined by risk factors. For evaluation of the general controls it is purposeful to select 5-10 samples of every examined process (e.g., giving IS user rights, management of IS changes etc.) and to recheck them. In some cases sample size may be increased or decreased.

• Evaluation of IS Internal Controls

Having performed analysis and testing of documents of IS general controls, an auditor evaluates status of internal control of information system. If sufficient internal control procedures are provided for and operate, and they are monitored, internal control may ensure information security (confidentiality, integrity, and accessibility)

Internal control may be evaluated using Capability Maturity Model

IS capability maturity may be identified evaluating all the IS internal control (giving one common score) or evaluating control of separate IS processes (giving scores for separate processes). Making final decision on reliability of the audited data, auditors have to evaluate the fact that data reliability may increase due to duplication of some data in paper versions.

Results of audits of IS general controls may be used when planning IS performance audits.

3.2.2 Audit of Application Controls

During the audit, it is usually sought to evaluate if the application controls is sufficient, if all the entries into information system are accurate, comprehensive, made in time, and people making and processing these entries are properly authorized.

Audit Resources of Application Controls

During the audits of application software control technical IS issues are usually not analysed, therefore such audits would have to be performed by general auditors. In cases when technical issues arise, IS auditors may be asked for help.

Performance of Application Controls Audit

Audit of application controls may be divided into three stages:

1. Analysis of documents of application controls;
2. Testing of application controls;
3. Evaluation of application controls.

• Analysis of Documents of Application Controls

In Auditing application controls the following aspects are taken into consideration:

• Organisation and documentation,
• Data entry,
• Data processing,
• Data transferring,
• Data output,
• Master data (data with little fluctuation, e.g., Litas/Euro exchange rate, VAT, depreciation standards).

Performing audit of application software it is recommended to use questionnaire presenting standard questions that may be freely chosen by an auditor.

• Testing of Application Controls

Evaluating application controls an auditor should perform 1-2 comprehensive monitoring of the system operation – from data entry to obtaining of result. Making a decision about reliability of this control, an auditor should make sure that the control was operating efficiently during all the audited period.

• Evaluation of Application Controls

Having performed analysis and testing procedures of application controls, an auditor estimates as to how reliable applications are. Results of audits of application controls may be used planning IS performance audits

3.2.3 Audit of IS Development Controls

During the audit of IS development controls it is examined whether IS development, installation, legitimating, and changes are properly controlled.

Audit of IS development controls may be also an object of IS general controls audit.

Audit of IS under development may violate independence of a public auditor, therefore in each individual case it has to be thought over whether it is worth conducting such audit. In order to mitigate this risk, it is suggested to limit oneself to evaluation of IS development controls, and not of the effectiveness and efficiency of the systems themselves.

Audit Performance

Performing audits of IS development controls the following aspects should be taken into consideration:

• Management of IS Development

Auditor has to evaluate IS design standards, programming standards, testing procedures, all the IS documentation, confirmation of system‘s users before implementation, incorporation of internal audit, division of duties between developers and operators.

Management of IS Changes

- Auditor has to evaluate monitoring and training, authorization for changes, as well as documentation, passwords security, making backups of IS data, physical protection of data, rotation of positions, testing procedures, and confirmation of changes.
- Management of IS implementation:
- Auditor has to evaluate IS testing and documentation, security of software and its carriers, division of duties among programmers, operators, and users.
- IS development controls in an audited entity may be also evaluated during audit of IS
- General controls.

3.3 Need for IS Audit Guideline

IS vs. Financial Audits

Unlike financial auditors, IS auditors require a different yardstick to measure materiality. Financial auditors ordinarily measure materiality in monetary terms, since what they audit is also measured and reported in monetary terms. IS auditors ordinarily perform audits of non-financial items, e.g., physical access controls, logical access controls, program change controls, and systems for personnel management, manufacturing control, design, quality control, password generation, credit card production and patient care. Therefore, IS auditors may need guidance on how materiality should be assessed to plan their audits effectively,

how to focus their effort on high-risk areas and how to assess the severity of any errors or weaknesses found.

This guideline provides guidance in applying IS auditing standards on audit materiality. The IS auditor should consider it in determining how to achieve implementation of the above standard, use professional judgment in its application and be prepared to justify any departure.

3.4 Information System Audit Planning

3.4.1 Assessing Materiality

The assessment of what is material is a matter of professional judgment and includes consideration of the effect and/or the potential effect on the organisation's ability to meet its business objectives in the event of errors, omissions, irregularities and illegal acts that may arise as a result of control weaknesses in the area being audited.

While assessing materiality, the IS auditor should consider:

- The aggregate level of error acceptable to management, the IS auditor, appropriate regulatory agencies and other stakeholders
- The potential for the cumulative effect of small errors or weaknesses to become material

To meet the audit objectives, the IS auditor should identify the relevant control objectives and, based on risk tolerance rate, determine what should be examined. With respect to a specific control objective, a material control is a control or group of controls without which control procedures do not provide reasonable assurance that the control objective will be met.

Where the IS audit objective relates to systems or operations that process financial transactions, the financial auditor's measure of materiality should be considered while conducting the IS audit.

The IS auditor should determine establishment of roles and responsibilities as well as a classification of information assets in terms of confidentiality, availability and integrity; access control rules on privileges management; and classification of information based upon degree of criticality and risk of exposure. Assessment should information include verification of:

- Information stored
- IS hardware
- IS architecture and software

• IS network infrastructure
• IS operations
• Development and test environment

The IS auditor should determine whether any IT general control deficiency could potentially become material. The significance of such deficient IT general controls should be evaluated in relation to their effect on application controls, i.e., whether the associated application controls are also ineffective. If the application deficiency is caused by the IT general control, then they are material. For example, if an application-based tax calculation is materially wrong and was caused by poor change controls to tax tables, then the application-based control (calculation) and the general control (changes) are materially weak.

The IS auditor should evaluate an IT general control's deficiency in relation to its effect on application controls and when aggregated against other control deficiencies. For example, a management decision not to correct an IT general control deficiency and its associated reflection on the control environment could become material when aggregated with other control deficiencies affecting the control environment.

The IS auditor should also note that failure to remediate a deficiency could become material.

The IS auditor should consider obtaining sign-off from appropriate stakeholders acknowledging they have disclosed existing material weakness that they are aware of in the organisation.

3.4.2 The Following are Examples of Measures that should be considered to Assess Materiality

• Criticality of the business processes supported by the system or operation
• Criticality of the information databases supported by the system or operation
• Number and type of application developed
• Number of users who use the information systems
• Number of managers and directors who work with the information systems classified by privileges
• Criticality of the network communications supported by the system or operation

- Cost of the system or operation (hardware, software, staff, third-party services, overheads or a combination of these)
- Potential cost of errors (possibly in terms of lost sales, warranty claims, irrecoverable development costs, cost of publicity required for warnings, rectification costs, health and safety costs, unnecessarily high costs of production, high wastage, etc.)
- Cost of loss of critical and vital information in terms of money and time to reproduce
- Effectiveness of countermeasures
- Number of accesses/transactions/inquiries processed per period
- Nature, timing and extent of reports prepared and files maintained
- Nature and quantities of materials handled (e.g., where inventory movements are recorded without values)
- Service level agreement requirements and cost of potential penalties
- Penalties for failure to comply with legal, regulatory and contractual requirements
- Penalties for failure to comply with public health and safety requirements

Control failures may potentially lead to monetary loss, competitive position, loss of trust or loss of reputation, apart from damaging the corporate image. The IS auditor should evaluate risks against possible countermeasures.

4.0 CONCLUSION

Controls are established to ensure that information systems work properly. Controls can be installed in the original system; the ISD, end-users, or others can add them once a system is in operation. Installing control is necessary but not sufficient. It is also necessary to answer questions such as the following: Are they effective? Did any breach of security occur? If so, what actions are required to prevent reoccurrence? These questions need to be answered by independent and unbiased observers. Such independent observers perform the information system auditing task.

An audit is an important part of any control system. In an organisational setting, it is usually referred to as a regular examination and check of financial and accounting records and procedures. Specially trained professionals who may be internal employees or external consultants execute auditing. In the information systems environment, auditing can be viewed as an additional layer of controls or safeguards.

5.0 SUMMARY

This unit is summarized as follows:

- Public auditors obtain part of the needed data from computerized accounting and Management systems. Auditors need to evaluate audit risk and reliability of the received information (audit evidence), therefore it is important to know how the audited entity information systems. controls
- IS audit is a process of evidence collection and evaluation allowing deciding whether a computer system (information system) ensures assets'security, data integrity, as well as helps to efficiently seek organisational goals and rationally use the resources.
- Audit is performed upon an assignment following the procedure established by the Auditor General or as a separate stage of financial or performance audit.
- Auditor should not consider computer processed and (or) transferred information of the audited entity reliable until he has proper supporting evidence. Such evidence could be obtained after getting assurance that internal control procedures of the system operate securely and properly.
- During the audit it is usually sought to evaluate if the application controls is sufficient, if all the entries into information system are accurate, comprehensive, made in time, and people making and processing these entries are properly authorized.
- Unlike financial auditors, IS auditors require a different yardstick to measure materiality.
- The assessment of what is material is a matter of professional judgment and includes consideration of the effect and/or the potential effect on the organisation's ability to meet its business objectives in the event of errors, omissions, irregularities and illegal acts that may arise as a result of control weaknesses in the area being audited.

6.0 TUTOR-MARKED ASSIGNMENT

1. Identify the objectives of Information systems audit.
2. Mention 5 aspects of auditing application control to be considered.

7.0 REFERENCES/FURTHER READINGS

G6 Materiality Concepts for Auditing Information Systems © 1999, 2008 ISACA.

Turban, E. McLean, E. and Wetherbe, J., (1999). Information *Technology Management. John Wiley & Sons Inc.*

UNIT 4 COPYRIGHT LAW AND ELECTRONIC ACCESS TO INFORMATION

CONTENTS

1.0 INTRODUCTION

Copyright is a legal concept, enacted by governments, giving the creator of an original work of authorship exclusive rights to control its distribution, usually for 70 years after the author's death, after which the work enters the public domain. Generally, it is “the right to copy”, but usually provides the author with other rights as well, such as the right to be credited for the work, to determine who may adapt the work to other forms, who may perform the work, who may financially benefit from it, and other, related rights. It is an intellectual property form (like the patent, the trademark, and the trade secret) applicable to any expressible form of an idea or information that is substantive and discrete. Copyright was initially conceived as a way for governments in Europe to restrict printing; the contemporary intent of copyright is to promote the creation of new works by giving authors control of and profit from them.

Copyright has been internationally standardized, lasting between fifty to a hundred years from the author's death, or a finite period for anonymous or corporate authorship; some jurisdictions have required formalities to establishing copyright, most recognize copyright in any completed work, without formal registration. Generally, copyright is enforced as a civil matter, though some jurisdictions do apply criminal sanctions.

Most jurisdictions recognize copyright limitations, allowing “fair” exceptions to the author’s exclusivity of copyright, and giving users certain rights. The development of the Internet, digital media, computer network technologies, such as peer-to-peer file sharing, have prompted reinterpretation of these exceptions, introduced new difficulties in enforcing copyright, and inspired additional challenges to copyright law's philosophic basis. Simultaneously, businesses with great economic dependence upon copyright have advocated the extension and expansion of their copy rights, and sought additional legal and technological enforcement.

2.0 OBJECTIVES

At the end of this unit you should be able to:

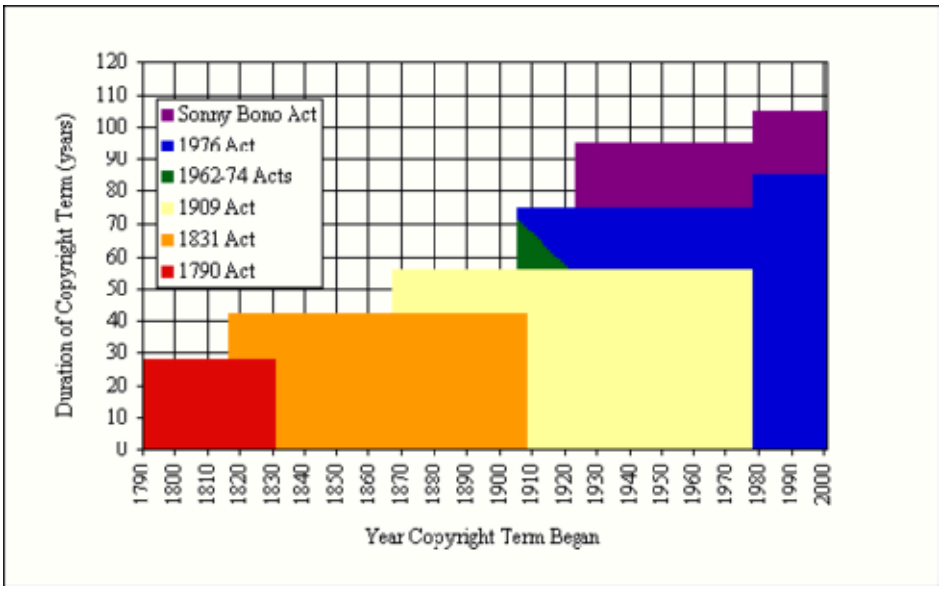
- define and know the concept of copyright
- trace the history and development of copyright
- answer the question of the scope and justification in applying copyright
- understand how to obtain and enforce copyright

• explain the limit and exceptions to copyright.

3.0 MAIN CONTENT

3.1 History

Figure 1: Expansion of U.S. Copyright Law



Copyright was invented after the advent of the printing press and with wider public literacy. As a legal concept, its origins in Britain were from a reaction to printers' monopolies at the beginning of the eighteenth century. Charles II of England was concerned with the unregulated copying of books and passed the Licensing Act of 1662 by Act of Parliament, which established a register of licensed books and required a copy to be deposited with the Stationer's Company, essentially continuing the licensing of material that had long been in effect.

The British Statute of Anne (1710) further alluded to individual rights of the author, beginning: "Whereas Printers, Booksellers, and other Persons, have of late frequently taken the Liberty of Printing... Books, and other Writings, without the Consent of the Authors... to their very great Detriment, and too often to the Ruin of them and their Families:..." A right to benefit financially from the work is articulated, and court rulings and legislation have recognized a right to control the work, such as ensuring that the integrity of it is preserved. An irrevocable right to be recognized as the work's creator appears in some countries' copyright laws.

The Statute of Anne was the first real copyright act, and gave the publishers the rights for a fixed period, after which the copyright expired. Copyright has grown from a legal concept regulating copying rights in the publishing of books and maps to one with a significant effect on nearly every modern industry, covering such items as sound recordings, films, photographs, software, and architectural works.

The Copyright Clause of the United States Constitution (1787) authorized copyright legislation: "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." That is, by guaranteeing them a period of time in which they alone could *profit from their works, they would be enabled and encouraged to invest* the time required to create them, and this would be good for society as a whole. A right to profit from the work has been the philosophical underpinning for much legislation extending the duration of copyright, to the life of the creator and beyond, to his heirs.

The 1886 Berne Convention first established recognition of copyrights among sovereign nations, rather than merely bilaterally. Under the Berne Convention, copyrights for creative works do not have to be asserted or declared, as they are automatically in force at creation: an author need not "register" or "apply for" a copyright in countries adhering to the Berne Convention. As soon as a work is "fixed", that is, written or recorded on some physical medium, its author is automatically entitled to all copyrights in the work, and to any derivative works unless and until the author explicitly disclaims them, or until the copyright expires. The Berne Convention also resulted in foreign authors being treated equivalently to domestic authors, in any country signed onto the Convention. The UK signed the Berne Convention in 1887 but did not implement large parts of it until 100 years later with the passage of the Copyright, Designs and Patents Act of 1988. The USA did not sign the Berne Convention until 1989.

The United States and most Latin American countries instead entered into the Buenos Aires Convention in 1910, which required a copyright notice (such as "all rights reserved") on the work, and signatory nations to limit the duration of copyrights to shorter renewable terms. The Universal Copyright Convention was drafted in 1952 as another less demanding alternative to the Berne Convention, and ratified by nations such as the Soviet Union and developing nations.

The regulations of the Berne Convention are incorporated into the World Trade Organisation's TRIPS agreement (1995), thus giving the Berne Convention effectively near-global application. The 2002 WIPO

Copyright Treaty enacted greater restrictions on the use of technology to copy works in the nations that ratified it.

3.2 Scope

Copyright may apply to a wide range of creative, intellectual, or artistic forms, or "works". Specifics vary by jurisdiction, but these can include poems, theses, plays, other literary works, movies, dances, musical compositions, audio recordings, paintings, drawings, sculptures, photographs, software, radio and television and broadcasts.

Copyright does not cover ideas and information themselves, only the form or manner in which they are expressed. For example, the copyright to a Mickey Mouse cartoon restricts others from making copies of the cartoon or creating derivative works based on Disney's particular anthropomorphic mouse, but doesn't prohibit the creation of other works about anthropomorphic mice in general, so long as they are different enough not to be judged copies of Disney's. In many jurisdictions, copyright law makes exceptions to these restrictions when the work is copied for the purpose of commentary or other related uses. Meanwhile, other laws may impose additional restrictions that copyright does not — such as trademarks and patents.

Copyright laws are standardized somewhat through international conventions such as the Berne Convention and Universal Copyright Convention. These multilateral treaties have been ratified by nearly all countries, and international organisations such as the European Union or World Trade Organisation require their member states to comply with them.

3.3 Justification

As with patents for physical objects, the granting of a copyright was ensured by governments to promote innovation and guarantee first-to-market protection for the owner of the copyright (historically, more likely the publisher than the creator). This government-sponsored monopoly thus provides innovation and general benefit to society as a whole, but allows for capitalistic pressures after the first-to-market advantage has been provided as a reward (and effort to cover R&D time for such works to be developed).

With the modern emergence of massive mass-media conglomerates however, the first-to-market advantage can be recouped within weeks instead of years. This point is highlighted easily by noting the millions of dollars investment in blockbuster movies are typically recouped within mere days, and the studios themselves even stop collecting ticket

sales income after typically one, though sometimes two weeks (which is when local theater owners finally start to collect revenue on ticket sales). Likewise, with the increasing use of technology such as Digital Rights Management to maintain studio control of content the time of monopolistic control of content is extended even beyond that guaranteed by law. This post-copyright restriction planning has come under fire as being disingenuous and even unethical use of the government awarded protection.

The solution to this criticism has been the heavy lobbying by Disney and artist unions to continually extend copyright protections, thus making DRM appear to be protecting copyrights that for all intents are effectively permanent... extending 25 years past the authors/artists death. The most recent extension of this corporate protection was provided by the bill sentimentally named the Sonny Bono Copyright Term Extension Act of 2000, which targeted Senator Bono's artistic heritage and recent death in an appeal to his colleges and the public support for such an act. Copyright lawyers commonly refer to this act as the Mickey Mouse **Protection Act due to the hundreds of millions of lobby dollars spent by** the Walt Disney Corporation to ensure its passing. Disney's vested interest in this act was due to the pending release of copyright on Steamboat Willie... the first Mickey Mouse cartoon whose success created the mega-cartoon corporation. Releasing Steamboat Willie to the public domain was seen as a slippery slope that Disney refused to allow due to their belief that copyrights should be indefinite and that they were entitled to society's granting of their monopoly.

Another widely debated issue is the relationship between copyrights and other forms of "intellectual property", and material property. Most scholars of copyright agree that it can be called a kind of property, because it involves the exclusion of others from something. But there is disagreement about the extent to which that fact should allow the transportation of other beliefs and intuitions about material possessions. This philosophical difference was highlighted by the Sony vs Disney case regarding record-able CDs and tape. At the time, Disney was attempting to ban VHS-recording machines as illegal devices attempting to impinge on their copyright. The United States Supreme Court disagreed and allowed the sale of VHS recording machines, and in a later, similar suit by Disney the US Supreme Court allowed the sale of recordable CDs and Mini-Discs. This repeated failure to gain government support of their position is what led Disney to try new tactics and lobby for increasing the length of copyright protection and eventually Digital Rights Management.

There are many other philosophical questions that arise in the jurisprudence of copyright. They include such problems as determining

when one work is “derived” from another, or deciding when information has been placed in a “tangible” or “material” form.

Some critics claim copyright law protects corporate interests while criminalizing legitimate use. Of particular concern is the increasing mound of orphaned works.

Orphaned works are those that were protected for so long that the original artist is no longer alive, and although the work may now be in the public domain, is no longer available due to physical decay of the paper, film, or physical form due to aging and lack of maintenance. The fact remains that less than 1% of all artistic works created in the United States belong to Disney or other corporations who will maintain their art for commercial gain. The bulk of artistic works do NOT generate any appreciable income after 5 years and due to copyright restrictions provide no motivation for museums, clearing houses, or enthusiast organisations to maintain records of the owner or a copy of the work. These orphaned works may not provide commercial benefit to the artists anymore, however they are fundamental to the fabric of society. As the orphan works disappear, historians lose valuable documents that hold insights into the evolution of phrases, social structure, and even the original source of new forms of art and genres that develop from them. Orphaned works are seen as justifiable losses to modern copyright lobbyists, equating them to an old chair or other form of property that has served its purpose and even if no longer economically viable, the copyright should be maintained in principle. This argument avoids the ethical implications of society losing the very art that it solicited by guaranteeing first-to-market rights.

3.4 Obtaining and Enforcing Copyright

Typically, a work must meet minimal standards of originality in order to qualify for copyright, and the copyright expires after a set period of time (some jurisdictions may allow this to be extended). Different countries impose different tests, although generally the requirements are low; in the United Kingdom there has to be some ‘'skill, labour and judgment,’ that has gone into it. In Australia and the United Kingdom it has been held that a single word is insufficient to comprise a copyright work. However, single words or a short string of words can sometimes be registered as a trademark instead.

Copyright law recognises the right of an author based on whether the work actually is an original creation, rather than based on whether it is unique; two authors may own copyright on two substantially identical works, if it is determined that the duplication was coincidental, and neither was copied from the other.

In all countries where the Berne Convention standards apply, copyright is automatic, and need not be obtained through official registration with any government office. Once an idea has been reduced to tangible form, for example by securing it in a fixed medium (such as a drawing, sheet music, photograph, a videotape, or a computer file), the copyright holder is entitled to enforce his or her exclusive rights. However, while registration isn't needed to exercise copyright, in jurisdictions where the laws provide for registration, it serves as prima facie evidence of a valid copyright and enables the copyright holder to seek statutory damages and attorney's fees. (In the USA, registering after an infringement only enables one to receive actual damages and lost profits.)

The original holder of the copyright may be the employer of the author rather than the author himself, if the work is a "work for hire". For example, in English law the Copyright, Designs and Patents Act 1988 provides that if a copyrighted work is made by an employee in the course of that employment, the copyright is automatically owned by the employer as a "Work for Hire."

Copyrights are generally enforced by the holder in a civil law court, but there are also criminal infringement statutes in some jurisdictions. While central registries are kept in some countries, which aid in proving claims of ownership, registering does not necessarily prove ownership, nor does the fact of copying (even without permission) necessarily prove that copyright was infringed. Criminal sanctions are generally aimed at serious counterfeiting activity, but are now becoming more commonplace as copyright collectives such as the RIAA are increasingly targeting the file sharing home Internet user. Thus far, however, most such cases against file sharers have been settled out of court

Copyright Notices in the U.S.

Prior to 1989, use of a copyright notice — consisting of the copyright symbol (©, the letter C inside a circle), the abbreviation "Copr.", or the word "Copyright", followed by the year of the first publication of the work and the name of the copyright holder — was part of United States statutory requirements. Several years may be noted if the work has gone through substantial revisions. The proper copyright notice for sound recordings of musical or other audio works is a sound recording copyright symbol (, the letter P inside a circle), which indicates a sound recording copyright. Similarly, the phrase. All rights reserved was once required to assert copyright.

In 1989, the U.S. enacted the Berne Convention Implementation Act, amending the 1976 Copyright Act to conform to most of the provisions of the Berne Convention. As a result, the use of copyright notices has become optional to claim copyright, because the Berne Convention makes copyright automatic. However, the lack of notice of copyright using these marks may have consequences in terms of reduced damages in an infringement lawsuit—using notices of this form may reduce the likelihood of a defense of "innocent infringement" being successful.

"Poor Man's Copyright"

A widely circulated strategy to avoid the cost of copyright registration is referred to as the "poor man's copyright." It proposes that the creator send the work to himself in a sealed envelope by registered mail, using the postmark to establish the date. This technique has not been recognized in any published opinions of the United States courts. The United States Copyright Office makes clear that the technique is no substitute for actual registration. The United Kingdom Intellectual Property Office discusses the technique but does not recommend its use.

3.5 Exclusive Rights

Several exclusive rights typically attach to the holder of a copyright:

- to produce copies or reproductions of the work and to sell those copies (mechanical rights; including, sometimes, electronic copies: distribution rights)
- to import or export the work
- to create derivative works (works that adapt the original work)
- to perform or display the work publicly (performance rights)
- to sell or assign these rights to others
- to transmit or display by radio or video (broadcasting rights)

The phrase "exclusive right" means that only the copyright holder is free to exercise those rights, and others are prohibited from using the work without the holders permission. Copyright is sometimes called a "negative right", as it serves to prohibit certain people (e.g., readers, viewers, or listeners, and primarily publishers and would be publishers) from doing something they would otherwise be able to do, rather than *permitting people (e.g., authors) to do something they would otherwise* be unable to do. In this way it is similar to the unregistered design right in English law and European law. The rights of the copyright holder also permit him/her to not use or exploit their copyright, for some or all of the term.

There is, however, a critique that rejects this assertion as being based on a philosophical interpretation of copyright law that is not universally shared. There is also debate on whether copyright should be considered a property right or a moral right. Many argue that copyright does not exist merely to restrict third parties from publishing ideas and information, and that defining copyright purely as a negative right is incompatible with the public policy objective of encouraging authors to create new works and enrich the public domain.

The right to adapt a work means to transform the way in which the work is expressed. Examples include developing a stage play or film script from a novel, translating a short story, and making a new arrangement of a musical work.

3.6 Limits and Exceptions to Copyright

Idea-Expression Dichotomy

Immanuel Kant in his 1785 essay Von der Unrechtmäßigkeit des *Büchernachdrucks distinguishes the physical from the ideational, the* thought involved from the book. This distinction is of critical importance to the near constant wrangling between publishers, other intermediaries, and the original, creative authors.

The First-Sale Doctrine and Exhaustion of Rights

Copyright law does not restrict the owner of a copy from reselling legitimately obtained copies of copyrighted works, provided that those copies were originally produced by or with the permission of the copyright holder. It is therefore legal, for example, to resell a copyrighted book or CD. In the United States this is known as the first-sale doctrine, and was established by the courts to clarify the legality of reselling books in second-hand bookstores. Some countries may have parallel importation restrictions that allow the copyright holder to control the aftermarket. This may mean for example that a copy of a book that does not infringe copyright in the country where it was printed **does infringe copyright in a country into which it is imported for** retailing. The first-sale doctrine is known as exhaustion of rights in other countries and is a principle that also applies, though somewhat differently, to patent and trademark rights. It is important to note that the first-sale doctrine permits the transfer of the particular legitimate copy involved. It does not permit making or distributing additional copies.

In addition, copyright, in most cases, does not prohibit one from acts such as modifying, defacing, or destroying his or her own legitimately obtained copy of a copyrighted work, so long as duplication is not involved. However, in countries that implement moral rights, a

copyright holder can in some cases successfully prevent the mutilation or destruction of a work that is publicly visible.

Fair Use and Fair Dealing

Copyright does not prohibit all copying or replication. In the United States, the fair use doctrine, codified by the Copyright Act of 1976 as 17 U.S.C. § 107, permits some copying and distribution without permission of the copyright holder or payment to same. The statute does not clearly define fair use, but instead gives four non-exclusive factors to consider in a fair use analysis. Those factors are:

- the purpose and character of the use;
- the nature of the copyrighted work;
- the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- the effect of the use upon the potential market for or value of the copyrighted work.

In the United Kingdom and many other Commonwealth countries, a similar notion of fair dealing was established by the courts or through legislation. The concept is sometimes not well defined; however in Canada, private copying for personal use has been expressly permitted by statute since 1999. In Australia, the fair dealing exceptions under the *Copyright Act 1968 (Cth) are a limited set of circumstances under which* copyrighted material can be legally copied or adapted without the copyright holder's consent. Fair dealing uses are research and study; review and critique; news reportage and the giving of professional advice (ie legal advice). Under current Australian law it is still a breach of copyright to copy, reproduce or adapt copyright material for personal or private use without permission from the copyright owner. Other technical exemptions from infringement may also apply, such as the temporary reproduction of a work in machine readable form for a computer.

In the United States the AHRA (Audio Home Recording Act Codified in Section 10, 1992) prohibits action against consumers making noncommercial recordings of music, in return for royalties on both media and devices plus mandatory copy-control mechanisms on recorders.

> *Section 1008. Prohibition on certain infringement actions*
> *No action may be brought under this title alleging infringement of copyright based on the manufacture, importation, or distribution of a digital audio recording device, a digital audio recording medium, an analog*

> *recording device, or an analog recording medium, or based on the noncommercial use by a consumer of such a device or medium for making digital musical recordings or analog musical recordings.*

Later acts amended US Copyright law so that for certain purposes 10 copies or more is construed to be commercial, but there is no general rule permitting such copying. Indeed making one complete copy of a work, or in many cases using a portion of it, for commercial purposes will not be considered fair use. The Digital Millennium Copyright Act prohibits the manufacture, importation, or distribution of devices whose intended use, or only significant commercial use, is to bypass an access or copy control put in place by a copyright owner. An appellate court has held that fair use is not a defense to engaging in such distribution.

Educational use is regarded as "fair use" in most jurisdictions, but the restrictions vary wildly from nation to nation.

Transfer and Licensing

A copyright, or aspects of it, may be assigned or transferred from one party to another. For example, a musician who records an album will often sign an agreement with a record company in which the musician agrees to transfer all copyright in the recordings in exchange for royalties and other considerations. The creator (and original copyright holder) benefits, or expects to, from production and marketing capabilities far beyond those of the author. In the digital age of music, music may be copied and distributed at minimal cost through the Internet, however the record industry attempts to provide promotion and marketing for the artist and his or her work so it can reach a much larger audience. A copyright holder need not transfer all rights completely, though many publishers will insist. Some of the rights may be transferred, or else the copyright holder may grant another party a non-exclusive license to copy and/or distribute the work in a particular region or for a specified period of time. A transfer or licence may have to meet particular formal requirements in order to be effective; see section 239 of the Australia Copyright Act 1968 (Cth). Under Australian law, it is not enough to pay for a work to be created in order to also own the copyright. The copyright itself must be expressly transferred in writing.

Under the U.S. Copyright Act, a transfer of ownership in copyright must be memorialized in a writing signed by the transferor. For that purpose, ownership in copyright includes exclusive licenses of rights. Thus exclusive licenses, to be effective, must be granted in a written

instrument signed by the grantor. No special form of transfer or grant is required. A simple document that identifies the work involved and the rights being granted is sufficient. Non-exclusive grants (often called non-exclusive licenses) need not be in writing under U.S. law. They can be oral or even implied by the behavior of the parties. Transfers of copyright ownership, including exclusive licenses, may and should be recorded in the U.S. Copyright Office. (Information on recording transfers is available on the Office's web site.) While recording is not required to make the grant effective, it offers important benefits, much like those obtained by recording a deed in a real estate transaction.

Copyright may also be licensed. Some jurisdictions may provide that certain classes of copyrighted works be made available under a prescribed statutory license (e.g. musical works in the United States used for radio broadcast or performance). This is also called a compulsory license, because under this scheme, anyone who wishes to copy a covered work does not need the permission of the copyright holder, but instead merely files the proper notice and pays a set fee established by statute (or by an agency decision under statutory guidance) for every copy made. Failure to follow the proper procedures would place the copier at risk of an infringement suit. Because of the difficulty of following every individual work, copyright collectives or collecting societies and performing rights organisations (such as ASCAP, BMI, and SESAC) have been formed to collect royalties for hundreds (thousands and more) works at once. Though this market solution bypasses the statutory license, the availability of the statutory fee still helps dictate the price per work collective rights organisations charge, driving it down to what avoidance of procedural hassle would justify.

Similar Legal Rights

Copyright law covers the creative or artistic expression of an idea. Patent law covers inventions. Trademark law covers distinctive terms, marks, and names that are used in relation to products or services as indicators of origin, as does (in a similar fashion), Trade dress. Registered designs law covers the look or appearance of a manufactured or functional article. Trade secret law covers secret or sensitive knowledge or information.

Although copyright and trademark laws are theoretically distinct, more than one type of them may cover the same item or subject matter. For example, in the case of the Mickey Mouse cartoon, the image and name of Mickey Mouse would be the subject of trademark legislation, while the cartoon itself would be subject to copyright. Titles and character

names from books or movies may also be trademarked while the works from which they are drawn may qualify for copyright.

Another point of distinction is that a copyright (and a patent) is generally subject to a statutorily-determined term, whereas a trademark registration may remain in force indefinitely if the trademark is periodically used and renewal fees continue to be duly paid to the relevant jurisdiction's trade marks office or registry. Once the term of a copyright has expired, the formerly copyrighted work enters the public domain and may be freely used or exploited by anyone. Courts in the United States and the United Kingdom have rejected the doctrine of a common law copyright. Public domain works should not be confused with works that are publicly available. Works posted in the internet for example, are publicly available, but are not generally in the public domain. Copying such works may therefore violate the author's copyright.

Useful Articles

If a pictorial, graphic or sculptural work is a useful article, it is copyrighted only if its aesthetic features are separable from its utilitarian features. A useful article is an article having an intrinsic utilitarian function that is not merely to portray the appearance of the article or to convey information. They must be separable from the functional aspect to be copyrighted.

There are two primary approaches to the separability issue: physical separability and conceptual separability. Physical separability is the ability to take the aesthetic thing away from the functional thing. Conceptual separability can be found in several different ways. It may be present if the useful article is also shown to be appreciated for its aesthetic appeal or by the design approach, which is the idea that separability is only available if the designer is able to make the aesthetic choices that are unaffected by the functional considerations. A question may also be asked of whether an individual would think of the aesthetic aspects of the work being separate from the functional aspects.

There are several different tests available for conceptual separability. The first, the Primary Use test, asks how the thing is primarily used: art or function. The second, the Marketable as Art test, asks can the article be sold as art, whether functional or not. This test does not have much backing, as almost anything can be sold as art. The third test, Temporal Displacement, asks could an individual conceptualize the article as art without conceptualizing functionality at the same time. Finally, the *Denicola test says that copyrightability should ultimately depend on the* extent to which the work reflects the artistic expression inhibited by

functional consideration. If something came to have a pleasing shape because there were functional considerations, the artistic aspect was constrained by those concerns.

Duration

Copyright subsists for a variety of lengths in different jurisdictions. The length of the term can depend on several factors, including the type of work (e.g. musical composition, novel), whether the work has been published or not, and whether the work was created by an individual or a corporation. In most of the world, the default length of copyright is the life of the author plus either 50 or 70 years. In the United States, the term for most existing works is a fixed number of years after the date of creation or publication. Under most countries' laws, copyrights expire at the end of the calendar year in question.

The length and requirements for copyright duration are subject to change by legislation, and since the early 20th century there have been a number of adjustments made in various countries, which can make determining the duration of a given copyright somewhat difficult. For example, the United States used to require copyrights to be renewed after 28 years to stay in force, and formerly required a copyright notice upon first publication to gain coverage. In Italy and France, there were post-wartime extensions that could increase the term by approximately 6 years in Italy and up to about 14 in France. Many countries have extended the length of their copyright terms (sometimes retroactively). International treaties establish minimum terms for copyrights, but individual countries may enforce longer terms than those.
In the United States, all books and other works published before 1923 have expired copyrights and are in the public domain. In addition, works published before 1964 that did not have their copyrights renewed 28 years after first publication year also are in the public domain, except that books originally published outside the US by non-Americans are exempt from this requirement, if they are still under copyright in their home country.

But if the intended exploitation of the work includes publication (or distribution of derivative work, such as a film based on a book protected by copyright) outside the U.S., the terms of copyright around the world must be considered. If the author has been dead more than 70 years, the work is in the public domain in most, but not all, countries. Some works are covered by copyright in Spain for 80 years after the author's death.

In 1998 the length of a copyright in the United States was increased by 20 years under the Copyright Term Extension Act. This legislation was strongly promoted by corporations that had valuable copyrights that

otherwise would have expired, and has been the subject of substantial criticism on this point.

As a curiosity, the famous work Peter Pan, or The Boy Who Wouldn't *Grow Up has a complex – and disputed – story of copyright expiry.*

Typefaces

In the United States, the Copyright Office maintains that typeface designs are not covered by copyright, and it will not accept applications for their registration. See 37. C.F.R. § 202.1(e). In Tufenkian *Import/Export Ventures, Inc. v. Einstein Moomjy, Inc., 338 F.3d 127,* 132 (2nd Cir. 2003), the United States Court of Appeals for the Second Circuit recognized this rule when it held, "the public domain includes, for example, both the generic shape of the letter 'L' and all of the elaborately more specific 'L's' from the hundreds of years of font designs that have fallen into the public domain." However, if a design is novel and "non-obvious," it may be covered by design patent. See, for example, U.S. Des. Patent No. 289,773, May 12, 1987), Charles Bigelow and Kris A. Holmes inventors. Germany (in 1981) passed a special extension (Schriftzeichengesetz) to the design patent law (Geschmacksmustergesetz) for protecting them. This permits typefaces being registered as designs in Germany, too. So far, the United States courts have not published any opinions discussing whether a computer program creating a particular font might be intellectual property protected by the copyright laws.

England recognized copyright in typeface at least as early as 1916. The current United Kingdom copyright statute, enacted in 1989, expressly refers to copyrights in typeface designs. The British law also applies to designs produced before 1989.

Accessible Copies

It is legal in several countries including the United Kingdom and the United States to produce alternative versions (for example, in large print or braille) of a copyrighted work to provide improved access to a work for blind and visually impaired persons without permission from the copyright holder.

3.7 Anti-Counterfeiting Trade Agreement (ACTA)

The Anti-Counterfeiting Trade Agreement (ACTA) is a proposed plurilateral trade agreement that would impose strict enforcement of intellectual property rights related to Internet activity and trade in information-based goods. The agreement is being secretly negotiated by

the governments of the United States, Japan, Switzerland, Australia, New Zealand, South Korea, Canada, and Mexico, and the European Commission. If adopted the treaty would establish an international coalition against copyright infringement, imposing strong, top-down enforcement of copyright laws in developed nations. The proposed agreement would allow border officials to search laptops, MP3 players, and cellular phones for copyright-infringing content. It would also impose new cooperation requirements upon Internet service providers (ISPs), including perfunctory disclosure of customer information, and restrict the use of online privacy tools. The proposal specifies a plan to encourage developing nations to accept the legal regime, as well.

4.0 CONCLUSION

Copyright is one of the oldest legislative schemes instituted to deal with crime, especially intellectual property crimes. However this legal cover is not often utilized, especially in developing countries, because of the seemingly delays of legal battles. So the challenge before copyright is more of awareness and implementation.

5.0 SUMMARY

- Copyright is a legal concept, enacted by governments, giving the creator of an original work of authorship exclusive rights to control its distribution, usually for 70 years after the author's death, after which the work enters the public domain.
- Copyright was invented after the advent of the printing press and with wider public literacy. As a legal concept, its origins in Britain were from a reaction to printers' monopolies at the beginning of the eighteenth century
- Copyright may apply to a wide range of creative, intellectual, or artistic forms, or "works". Specifics vary by jurisdiction, but these can include poems, theses, plays, other literary works, movies, dances, musical compositions, audio recordings, paintings, drawings, sculptures, photographs, software, radio and television and broadcasts.
- As with patents for physical objects, the granting of a copyright was ensured by governments to promote innovation and guarantee first-to-market protection for the owner of the copyright (historically, more likely the publisher than the creator).
- Typically, a work must meet minimal standards of originality in order to qualify for copyright, and the copyright expires after a set period of time (some jurisdictions may allow this to be extended).
- There are sseveral exclusive rights typically attach to the holder of a copyright:

- Copyright law does not restrict the owner of a copy from reselling legitimately obtained copies of copyrighted works, provided that those copies were originally produced by or with the permission of the copyright holder.
- The Anti-Counterfeiting Trade Agreement (ACTA) is a proposed plurilateral trade agreement that would impose strict enforcement of intellectual property rights related to Internet activity and trade in information-based goods.

6.0 TUTOR-MARKED ASSIGNMENT

1. Discuss briefly the scope of copyright
2. Identify 4 non-exclusive factors to consider in a fair deal analysis

7.0 REFERENCES/FURTHER READING

Patterson, (1968). Copyright in Historical Perspective. Vanderbilt Univ. Press.

Express Newspaper Plc v News (UK) Plc, F.S.R. 36 (1991).

Copyright Act of 1976, Pub.L. 94-553, 90 Stat. 2541, § 401(a) (October 19, 1976)

U.S. Copyright Office - Information Circular
Copyright in General: I've heard about a "poor man's copyright". What is it?, U.S Copyright Office

Copyright Registers', United Kingdom Intellectual Property Office

International comparison of Educational "fair use" legislation
U.S. Copyright Office - Copyright Law: Chapter 1

Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc., 338 F. 3d 127, 132 (2nd Cir. 2003) (via FindLaw)

"Stephenson, Blake and Co. v. Grant, Legros & Co., 115 L.T.R. 666, 61 Sol. J. 55 (1916), Reprinted in E.J. MacGillivray, Copyright Cases 1911-1916 326-329 (1969), aff'd 116 L.T.R. 268 (1917),

Copyright, Designs and Patents Act, 1988, ch. 48, § 54 (England)

Copyright (Visually Impaired Persons) Act 2002 (England):

Geiger, Andrea (2008-04-30). "A View From Europe: The High Price of Counterfeiting, and Getting Real about Enforcement, The Hill. Retrieved on 2008-05-27.

Pilieci, Vito "Copyright Deal Could Toughen Rules Governing info on iPods, Computers", Vancouver Sun. Retrieved on 2008-05-27.

"Proposed US ACTA Multi-Lateral Intellectual Property Trade Agreement (2007)". Wikileaks (May 22, 2008).

Jason Mick "Wikileaks Airs U.S. Plans to Kill Pirate Bay, Monitor ISPs With Multinational ACTA Proposal". DailyTech. (May 23, 2008).

Weeks, Carly "Anti-Piracy Strategy will Help Government to spy, Critic says", The Globe and Mail. Retrieved on 2008-05-27.

UNIT 5 INTERNET FIREWALL

CONTENTS

1.0 INTRODUCTION

Security has become one of the primary concerns when an organisation connects its private network to the Internet. Regardless of the business, an increasing number of users on private networks are demanding access to Internet services such as the World Wide Web (WWW), Internet mail, Telnet, and File Transfer Protocol (FTP). In addition, corporations want to offer WWW home pages and FTP servers for public access on the Internet.

Network administrators have increasing concerns about the security of their networks when they expose their organisation's private data and networking infrastructure to Internet crackers. To provide the required level of protection, an organisation needs a security policy to prevent unauthorized users from accessing resources on the private network and to protect against the unauthorized export of private information. Even if an organisation is not connected to the Internet, it may still want to establish an internal security policy to manage user access to portions of the network and protect sensitive or secret information.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the concerns of network security administrators
- define Internet firewalls
- identify the benefits and limitations associated with Internet firewalls
- identify the tools used by hackers as way of knowing how to counter their operations
- identify the kind of decisions to make in then design of firewalls
- identify and differentiate the types of firewalls.

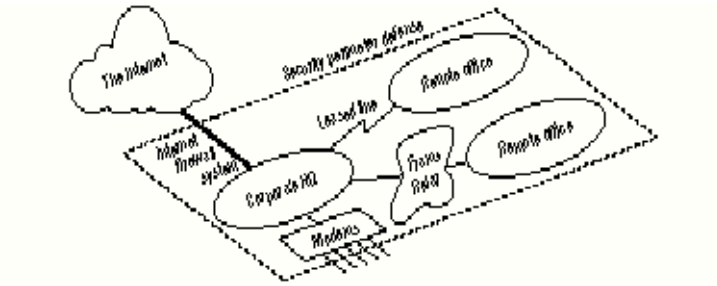


Figure 1. Security Policy Creates a Perimeter Defense

3.0 MAIN CONTENT

3.1 Internet Firewalls

An Internet firewall is a system or group of systems that enforces a security policy between an organisation's network and the Internet. The firewall determines which inside services may be accessed from the outside, which outsiders are permitted access to the permitted inside services, and which outside services may be accessed by insiders. For a firewall to be effective, all traffic to and from the Internet must pass through the firewall, where it can be inspected (Figure 1). The firewall must permit only authorized traffic to pass, and the firewall itself must be immune to penetration. Unfortunately, a firewall system cannot offer any protection once an attacker has gotten through or around the firewall.

It is important to note that an Internet firewall is not just a router, a bastion host, or a combination of devices that provides security for a network. The firewall is part of an overall security policy that creates a perimeter defense designed to protect the information resources of the organisation. This security policy must include published security guidelines to inform users of their responsibilities; corporate policies defining network access, service access, local and remote user authentication, dial-in and dial-out, disk and data encryption, and virus protection measures; and employee training. All potential points of network attack must be protected with the same level of network security. Setting up an Internet firewall without a comprehensive security policy is like placing a steel door on a tent.

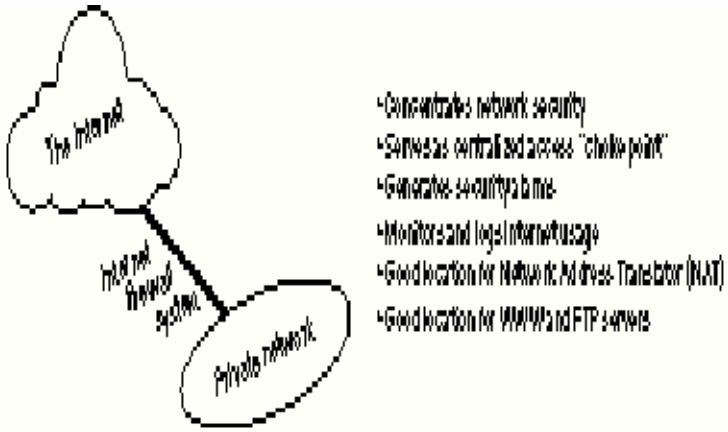


Figure 2. Benefits of an Internet Firewall

Benefits of an Internet Firewall

Internet firewalls manage access between the Internet and an organisation's private network (Figure 2). Without a firewall, each host system on the private network is exposed to attacks from other hosts on the Internet. This means that the security of the private network would depend on the “hardness” of each host's security features and would be only as secure as the weakest system.

Internet firewalls allow the network administrator to define a centralized "choke point" that keeps unauthorized users such as hackers, crackers, vandals, and spies out of the protected network; prohibits potentially vulnerable services from entering or leaving the protected network; and provides protection from various types of routing attacks. An Internet firewall simplifies security management, since network security is consolidated on the firewall systems rather than being distributed to every host in the entire private network.

Firewalls offer a convenient point where Internet security can be monitored and alarms generated. It should be noted that for organisations that have connections to the Internet, the question is not whether but when attacks will occur. Network administrators must audit and log all significant traffic through the firewall. If the network administrator doesn't take the time to respond to each alarm and examine logs on a regular basis, there is no need for the firewall, since the network administrator will never know if the firewall has been successfully attacked.

For the past few years, the Internet has been experiencing an address space crisis that has made registered IP addresses a less precious commodity. This means that organisations wanting to connect to the Internet may not be able to obtain enough registered IP addresses to meet the demands of their user population. An Internet firewall is a logical place to deploy a Network Address Translator (NAT) that can help alleviate the address space shortage and eliminate the need to renumber when an organisation changes Internet service providers (ISPs).

An Internet firewall is the perfect point to audit or log Internet usage. This permits the network administrator to justify the expense of the Internet connection to management, pinpoint potential bandwidth bottlenecks, and provide a method for departmental charge-backs if this fits the organisation's financial model.

An Internet firewall can also offer a central point of contact for information delivery service to customers. The Internet firewall is the ideal location for deploying World Wide Web and FTP servers. The firewall can be configured to allow Internet access to these services, while prohibiting external access to other systems on the protected network.

Finally, some might argue that the deployment of an Internet firewall creates a single point of failure. It should be emphasized that if the connection to the Internet fails, the organisation's private network will

still continue to operate--only Internet access is lost. If there are multiple points of access, each one becomes a potential point of attack that the network administrator must firewall and monitor regularly.

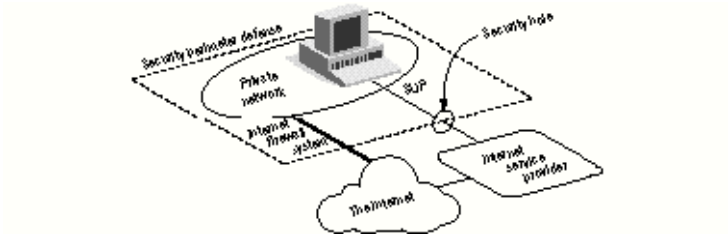


Figure 3. A Connection Circumventing an Internet Firewall

Limitations of an Internet Firewall

An Internet firewall cannot protect against attacks that do not go through the firewall. For example, if unrestricted dial-out is permitted from inside the protected network, internal users can make a direct SLIP or PPP connection to the Internet. Savvy users who become irritated with the additional authentication required by firewall proxy servers may be tempted to circumvent the security system by purchasing a direct SLIP or PPP connection to an ISP. Since these types of connections bypass the security provided by the most carefully constructed firewall, they create a significant potential for back-door attacks (Figure 3). Users must be made aware that these types of connections are not permitted as part of the organisation's overall security architecture.
Internet firewalls cannot protect against the types of threats posed by traitors or unwitting users. Firewalls do not prohibit traitors or corporate spies from copying sensitive data onto floppy disks or PCMCIA cards and removing them from a building. Firewalls do not protect against attacks where a hacker, pretending to be a supervisor or a befuddled new employee, persuades a less sophisticated user into revealing a password or granting them “temporary” network access. Employees must be educated about the various types of attacks and about the need to guard and periodically change their passwords.

Internet firewalls cannot protect against the transfer of virus-infected software or files. Since there are so many different viruses, operating systems, and ways of encoding and compressing binary files, an Internet firewall cannot be expected to accurately scan each and every file for potential viruses. Concerned organisations should deploy anti-viral software at each desktop to protect against their arrival from floppy disks or any other source.

Finally, Internet firewalls cannot protect against data-driven attacks. A data-driven attack occurs when seemingly harmless data is mailed or copied to an internal host and is executed to launch an attack. For

example, a data-driven attack could cause a host to modify security-related files, making it easier for an intruder to gain access to the system. As we will see, the deployment of proxy servers on a bastion host is an excellent means of prohibiting direct connections from the outside and reducing the threat of data-driven attacks.

3.2 The Hacker's Toolbox

It is difficult to describe a typical hacker attack because intruders have different levels of technical expertise and many different motivations. Some hackers are intrigued by the challenge, others just want to make life more difficult for others, and still others are out to steal sensitive data for profit.

Information Gathering

Generally, the first step in a break-in is some form of information gathering. The goal is to construct a database of the target organisation's network and gather information about the hosts residing on each of the networks. There are a number of tools that a hacker can use to collect this information:

- The SNMP protocol can be used to examine the routing table of an unsecured router to learn intimate details about the target organisation's network topology.
- The Trace Route program can reveal intermediate network numbers and routers in the path to a specific host.
- The Who is protocol is an information service that can provide data about all DNS domains and the system administrators responsible for each domain. However, this information is usually out of date.
- DNS servers can access a list of host IP addresses and their corresponding host names.
- The Finger protocol can reveal detailed information about the users (login names, phone numbers, time they last logged in, etc.) of a specified host.
- The Ping program can be employed to locate a particular host and determine its reachability. This simple tool can be used in a short scanning program that pings every possible host address on a network to construct a list of the hosts actually residing on the network.

Probing Systems for Security Weaknesses

After information about the targeted organisation's network is gathered, the hacker attempts to probe each host for security weaknesses. There

are a number of tools that a hacker can use to automatically scan the individual hosts residing on a network; for example:

- Since the list of known service vulnerabilities is rather short, a knowledgeable hacker can write a small program that attempts to connect to specific service ports on a targeted host. The output of the program is a list of hosts that support services that are exposed to attack.
- There are several publicly available tools, such as the Internet Security Scanner (ISS) or the Security Analysis Tool for Auditing Networks (SATAN), that scan an entire domain or subnetwork and look for security holes. These programs determine the weaknesses of each system with respect to several common system vulnerabilities. Intruders use the information collected from these scans to gain unauthorized access to the targeted organisation's systems.

A clever network administrator can use these tools within their private network to discover potential security weaknesses and determine which hosts need to be updated with new software patches.

Accessing Protected Systems

The intruder uses the results of the host probes to target a specific system for attack. After gaining access to a protected system, the hacker has many options available:

- The intruder can attempt to destroy evidence of the assault and open new security holes or back doors in the compromised system in order to have continued access even if the original attack is discovered.
- The intruder can install packet sniffers that include Trojan horse binaries that hide the sniffing activity on the installed systems. The packet sniffers collect account names and passwords for Telnet and FTP services that allow the hacker to spread the attack to other machines.
- The intruder can find other hosts that trust the compromised system. This allows the hacker to exploit the vulnerabilities of a single host and spread the attack across the entire organisation's network.
- If the hacker can obtain privileged access on a compromised system, he or she can read mail, search private files, steal private files, and destroy or corrupt important data.

3.3 Basic Firewalls Design Decisions

When designing an Internet firewall, there are a number of decisions that must be addressed by the network administrator:

- The stance of the firewall
- The overall security policy of the organisation
- The financial cost of the firewall
- The components or building blocks of the firewall system

Stance of the Firewall

The stance of a firewall system describes the fundamental philosophy of the organisation. An Internet firewall may take one of two diametrically opposed stances:

- Everything not Specifically Permitted is Denied. This stance assumes that a firewall should block all traffic, and that each desired service or application should be implemented on a case-by-case basis. This is the recommended approach. It creates a very secure environment, since only carefully selected services are supported. The disadvantage is that it places security ahead of ease of use, limiting the number of options available to the user community.

- Everything not Specifically Denied is permitted. This stance assumes that a firewall should forward all traffic, and that potentially harmful service should be shut off on a case-by-case basis. This approach creates a more flexible environment, with more services available to the user community. The disadvantage is that it puts ease of use ahead of security, putting the network administrator in a reactive mode and making it increasingly difficult to provide security as the size of the protected network grows.

Security Policy of the Organisation

As discussed earlier, an Internet firewall does not stand alone--it is part of the organisation's overall security policy, which defines all aspects of its perimeter defense. To be successful, organisations must know what they are protecting. The security policy must be based on a carefully conducted security analysis, risk assessment, and business needs analysis. If an organisation does not have a detailed security policy, the most carefully crafted firewall can be circumvented to expose the entire private network to attack.

Cost of the Firewall

How much security can the organisation afford? A simple packet-filtering firewall can have a minimal cost since the organisation needs a router to connect to the Internet, and packet filtering is included as part of the standard router feature set. A commercial firewall system provides increased security but may cost from U.S.$4,000 to $30,000, depending on its complexity and the number of systems protected. If an organisation has the in-house expertise, a home-brewed firewall can be constructed from public domain software, but there are still costs in terms of the time to develop and deploy the firewall system. Finally, all firewalls require continuing support for administration, general maintenance, software updates, security patches, and incident handling.

Components of the Firewall System

After making decisions about firewall stance, security policy, and budget issues, the organisation can determine the specific components of its firewall system. A typical firewall is composed of one or more of the following building blocks:

- Packet-filtering router
- Application-level gateway (or proxy server)
- Circuit-level gateway

The remainder of this paper discusses each of these building blocks and describes how they can work together to build an effective Internet firewall system.

3.4 Types of Firewalls

Firewall Example #1: Packet-Filtering Router
The most common Internet firewall system consists of nothing more than a packet-filtering router deployed between the private network and the Internet (Figure 4). A packet-filtering router performs the typical routing functions of forwarding traffic between networks as well as using packet-filtering rules to permit or deny traffic. Typically, the filter rules are defined so that hosts on the private network have direct access to the Internet, while hosts on the Internet have limited access to systems on the private network. The external stance of this type of firewall system is usually that everything not specifically permitted is denied.

Figure 4. Packet-Filtering Router Firewall

Although this firewall system has the benefit of being inexpensive and transparent to users, it possesses all of the limitations of a packet-filtering router such as exposure to attacks from improperly configured filters and attacks that are tunneled over permitted services. Since the direct exchange of packets is permitted between outside systems and inside systems, the potential extent of an attack is determined by the total number of hosts and services to which the packet-filtering router permits traffic. This means that each host directly accessible from the Internet needs to support sophisticated user authentication and needs to be regularly examined by the network administrator for signs of attack. Also, if the single packet-filtering router is penetrated, every system on the private network may be compromised.

Firewall Example #2: Screened Host Firewall

The second firewall example employs both a packet-filtering router and a bastion host (Figure 9). This firewall system provides a higher level of security than the previous example because it implements both network-layer security (packet-filtering) and application-layer security (proxy services). Also, an intruder has to penetrate two separate systems before the security of the private network can be compromised.

For this firewall system, the bastion host is configured on the private network with a packet-filtering router between the Internet and the bastion host. The filtering rules on the exposed router are configured so that outside systems can access only the bastion host; traffic addressed to all other internal systems is blocked. Since the inside hosts reside on the same network as the bastion host, the security policy of the organisation determines whether inside systems are permitted direct access to the Internet, or whether they are required to use the proxy services on the bastion host. Inside users can be forced to use the proxy services by configuring the router's filter rules to accept only internal traffic originating from the bastion host.

One of the benefits of this firewall system is that a public information server providing Web and FTP services can be placed on the segment shared by the packet-filtering router and the bastion host. If the strongest security is required, the bastion host can run proxy services that require both internal and external users to access the bastion host before communicating with the information server. If a lower level of security is adequate, the router may be configured to allow outside users direct access to the public information server.

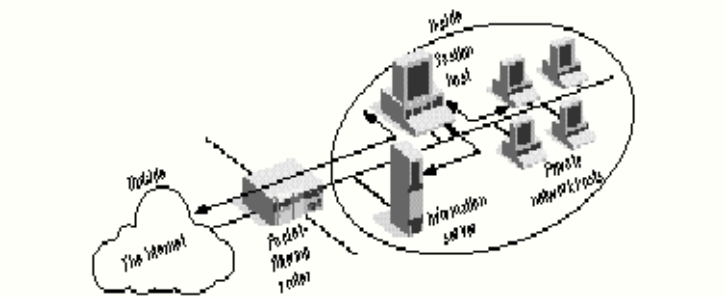


Figure 5. Screened Host Firewall System (Single-Homed Bastion Host)

An even more secure firewall system can be constructed using a dual-homed bastion host system (Figure 5). A dual-homed bastion host has two network interfaces, but the host's ability to directly forward traffic between the two interfaces bypassing the proxy services is disabled. The physical topology forces all traffic destined for the private network through the bastion host and provides additional security if outside users are granted direct access to the information server.

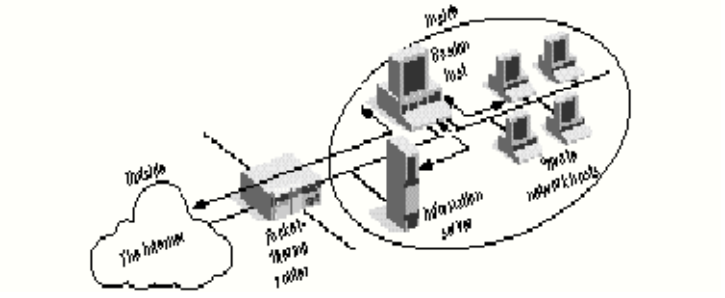


Figure 6. Screened Host Firewall System (Dual-Homed Bastion Host)

Since the bastion host is the only internal system that can be directly accessed from the Internet, the potential set of systems open to attack is limited to the bastion host. However, if users are allowed to log on to the bastion host, the potential set of threatened systems expands to include the entire private network, since it is much easier for an intruder to compromise the bastion host if they are allowed to log on. It is critical

that the bastion host be hardened and protected from penetration and that users never be allowed to log on to the bastion host.

Firewall Example #3: "Demilitarized Zone" or Screened-Subnet Firewall

The final firewall example employs two packet-filtering routers and a bastion host (Figure 7). This firewall system creates the most secure firewall system, since it supports both network- and application-layer security while defining a “demilitarized zone” (DMZ) network. The network administrator places the bastion host, information servers, modem pools, and other public servers on the DMZ network. The DMZ network functions as a small, isolated network positioned between the Internet and the private network. Typically, the DMZ is configured so that systems on the Internet and systems on the private network can access only a limited number of systems on the DMZ network, but the direct transmission of traffic across the DMZ network is prohibited.

For incoming traffic, the outside router protects against the standard external attacks (source IP address spoofing, source routing attacks, etc.) and manages Internet access to the DMZ network. It permits external systems to access only the bastion host (and possibly the information server). The inside router provides a second line of defense, managing DMZ access to the private network by accepting only traffic originating from the bastion host.

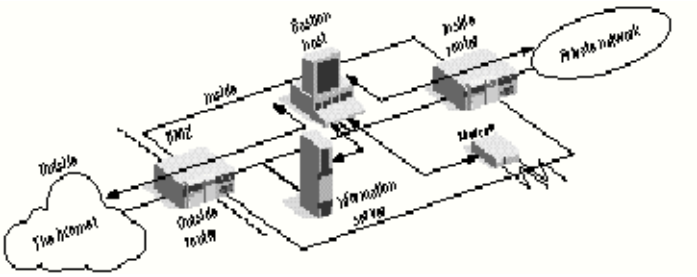


Figure 7. Screened-Subnet Firewall System

For Internet-bound traffic, the inside router manages private network access to the DMZ network. It permits internal systems to access only the bastion host (and possibly the information server). The filtering rules on the outside router require use of the proxy services by accepting only Internet-bound traffic from the bastion host.

There are several key benefits to the deployment of a screened subnet firewall system:

- An intruder must crack three separate devices (without detection) to infiltrate the private network: the outside router, the bastion host, and the inside router.
- Since the outside router advertises the DMZ network only to the Internet, systems on the Internet do not have routes to the protected private network. This allows the network manager to ensure that the private network is "invisible," and that only selected systems on the DMZ are known to the Internet via routing table and DNS information exchanges.
- Since the inside router advertises the DMZ network only to the private network, systems on the private network do not have direct routes to the Internet. This guarantees that inside users must access the Internet via the proxy services residing on the bastion host.
- Packet-filtering routers direct traffic to specific systems on the DMZ network, eliminating the need for the bastion host to be dual-homed.
- The inside router supports greater packet throughput than a dual-homed bastion host when it functions as the final firewall system between the private network and the Internet.
- Since the DMZ network is a different network than the private network, a Network Address Translator (NAT) can be installed on the bastion host to eliminate the need to renumber or resubnet the private network.

Electronic Signature

The digital signature algorithm (DSA) is another form of firewall. It is a digital signature and verification mechanism used for digital, rather than written signature. DSA enables the verification of signature, message origin, and message integrity without giving away information that would make signature forgery possible. DSA achieves this by allotting two different digital keys to each signature bearer a secret private key for encrypting the message and a public key for decrypting it. Only the signature bearer knows this private key, while the entire network user knows the public key. The details of digital signature will be discussed in details in Unit

4.0 CONCLUSION

There is no single correct answer for the design and deployment of Internet firewalls. Many different factors such as their corporate security policy, the technical background of their staff, cost, and the perceived threat of attack will influence each organisation's decision. This paper focused on many of the issues relating to the construction of Internet firewalls, including their benefits, limitations, building blocks, and examples of firewall system topologies. Since the benefits of connecting to the global Internet probably exceed its costs, network managers

should proceed with an awareness of the dangers and an understanding that, with the proper precautions, their networks can be as safe as they need them to be.

5.0 SUMMARY

- Security has become one of the primary concerns when an organisation connects its private network to the Internet. Regardless of the business, an increasing number of users on private networks are demanding access to Internet services such as the World Wide Web (WWW), Internet mail, Telnet, and File Transfer Protocol (FTP).
- An Internet firewall is a system or group of systems that enforces a security policy between an organisation's network and the Internet. The firewall determines which inside services may be accessed from the outside, which outsiders are permitted access to the permitted inside services, and which outside services may be accessed by insiders.
- It is difficult to describe a typical hacker attack because intruders have different levels of technical expertise and many different motivations. Some hackers are intrigued by the challenge, others just want to make life more difficult for others, and still others are out to steal sensitive data for profit.
- After information about the targeted organisation's network is gathered, the hacker attempts to probe each host for security weaknesses.
- When designing an Internet firewall, there are a number of decisions that must be addressed by the network administrator
- As discussed earlier, an Internet firewall does not stand alone--it is part of the organisation's overall security policy, which defines all aspects of its perimeter defense.
- The most common Internet firewall system consists of nothing more than a packet-filtering router deployed between the private network and the Internet

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention firewall basic design consideration in decision.
2. What are the basic components of the firewall system?

7.0 REFERENCES/FURTHER READINGS

3Com Corporation, NASDAQ.COMS, 1996.

MODULE 3

Unit 1 Digital Signature and Electronic Signature
Unit 2 Biometric Identification
Unit 3 Fraud Prevention
Unit 4 Sanctions against Plastic Card Fraud: The Case of Australia

UNIT 1 DIGITAL SIGNATURE AND ELECTRONIC SIGNATURE

CONTENTS

1.0 INTRODUCTION

A digital signature or digital signature scheme is a type of asymmetric cryptography used to simulate the security properties of a handwritten signature on paper. Digital signature schemes consist of at least three algorithms: a key generation algorithm, a signature algorithm, and a verification algorithm. A digital signature mainly provides authentication of a "message". In theory it can also provide repu- diation, meaning that the authenticity of signed messages can be publicly verified, not only by the intended recipient. Messages may be anything, from electronic mail to a contract, or even a message sent in a more complicated cryptographic protocol.

Digital signatures are often used to implement electronic signatures, a broader term that refers to any electronic data that carries the intent of a signature, but not all electronic signatures use digital signatures. In some countries, including the United States, and in the European Union, electronic signatures have legal significance. However, laws concerning electronic signatures do not always make clear their applicability towards cryptographic digital signatures, leaving their legal importance somewhat unspecified.

Digital signature is a subset of electronic signature. It is an emerging form of electronic signature.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define digital and electronic signatures
- compare and contrast digital and electronic signatures
- identify the basic features of digital and electronic signatures
- explain the applications of digital and electronic signatures
- trace the history and development of digital and electronic signatures
- identify the drawbacks associated with digital and electronic signatures.

3.0 MAIN CONTENT

3.1 Definition

A digital signature scheme typically consists of three algorithms:

- A key generation algorithm that selects a private key uniformly at random from a set of possible private keys. The algorithm outputs the private key and a corresponding public key.
- A signing algorithm which, given a message and a private key, produces a signature.
- A signature verifying algorithm which given a message, public key and a signature, either accepts or rejects.

Two main properties are required. First, a signature generated from a fixed message and fixed private key should verify on that message and the corresponding public key. Secondly, it should be computationally infeasible to generate a valid signature for a party who does not possess the private key.

3.2 History

In the famous paper "New Directions in Cryptography", Whitfield Diffie and Martin Hellman first described the notion of a digital signature scheme, although they only conjectured that such schemes existed. Soon afterwards, Ronald Rivest, Adi Shamir, and Len Adleman invented the RSA algorithm that could be used for primitive digital signatures (Note that this just serves as a proof-of-concept, and "plain" RSA signatures are not secure.) The first widely marketed software package to offer digital signature was Lotus Notes 1.0, released in 1989, which used the RSA algorithm.

Basic RSA signatures are computed as follows. To generate RSA signature keys, one simply generates an RSA key pair containing a modulus N that is the product of two large primes, along with integers e and d such that e d = 1 mod f(N), where f is the Euler phi-function. The signer's public key consists of N and e, and the signer's secret key contains

To sign a message m, the signer computes s=md mod N. To verify, the receiver checks that se = m mod N.

As noted earlier, this basic scheme is not very secure. To prevent attacks, one can first apply a cryptographic hash function to the message *m and then apply the RSA algorithm described above to the result. This* approach can be proven secure in the so-called random oracle model.

Other digital signature schemes were soon developed after RSA, the earliest being Lamport signatures, Merkle signatures (also known as "Merkle trees" or simply Hash trees"), and Rabin signatures.

In 1984, Shafi Goldwasser, Silvio Micali, and Ronald Rivest became the first to rigorously define the security requirements of digital signature schemes. They described a hierarchy of attack models for signatures, and also present the GMR signature scheme, the first that can be proven to prevent even an existential forgery against a chosen-message attack.

Most early signature schemes were of a similar type: they involve the use of a trapdoor permutation, such as the RSA function, or in the case of the Rabin signature scheme, computing square modulo composite n. A trapdoor permutation family is a family of permutations, specified by a parameter that is easy to compute in the forward direction, but is difficult to compute in the reverse direction. However, for every parameter there is a "trapdoor" that enables easy computation of the reverse direction. Trapdoor permutations can be viewed as public-key encryption systems, where the parameter is the public key and the trapdoor is the secret key, and where encrypting corresponds to computing the forward direction of the permutation, while decrypting corresponds to the reverse direction. Trapdoor permutations can also be viewed as digital signature schemes, where computing the reverse direction with the secret key is thought of as signing, and computing the forward direction is done to verify signatures. Because of this correspondence, digital signatures are often described as based on public-key cryptosystems, where signing is equivalent to decryption and verification is equivalent to encryption, but this is not the only way digital signatures are computed.

Used directly, this type of signature scheme is vulnerable to a key-only existential forgery attack. To create a forgery, the attacker picks a random signature s and uses the verification procedure to determine the message m corresponding to that signature. In practice, however, this type of signature is not used directly, but rather, the message to be signed is first hashed to produce a short digest that is then signed. This forgery attack, then, only produces the hash function output that corresponds to s, but not a message that leads to that value, which does not lead to an attack. In the random oracle model, this hash-and-decrypt form of signature is existentially unforgeable, even against a chosen-message attack.

There are several reasons to sign such a hash (or message digest) instead of the whole document.

- For Efficiency: The signature will be much shorter and thus save time since hashing is generally much faster than signing in practice.
- For Compatibility: Messages are typically bit strings, but some signature schemes operate on other domains (such as, in the case of RSA, numbers modulo a composite number N). A hash function can be used to convert an arbitrary input into the proper format.
- For Integrity: Without the hash function, the text "to be signed" may have to be split (separated) in blocks small enough for the signature scheme to act on them directly. However, the receiver of the signed blocks is not able to recognize if all the blocks are present and in the appropriate order.

3.3 Digital Signature vs Electronic Signature

The term electronic signature has several meanings. Among the more expansive is that given by US law, influenced by ABA committee white papers and the uniform law promulgated by the National Conference of Commissioners on Uniform State Laws (NCCUSL). Under the Uniform Electronic Transactions Act or “UETA” released by NCCUSL in 1999., the term means “an electronic sound, symbol, or process, attached to or logically associated with a record and executed or adopted by a person with the intent to sign the record.” This definition and many other core concepts of UETA are echoed in the U.S. ESign Act of 2000 46 US states, the District of Columbia, and the US Virgin Islands have enacted UETA.

The concept itself is not new. US and other common law contain references to telegraph signatures and faxed signatures, some as far back as the mid-19th century. For that matter, the text of, and comments to, US Federal Rules of Evidence 1001, 1002, and 1003, among others, give good support for the proposition that electronic records and signatures would be admissible in court.

There is confusion between the terms electronic signature and digital signature. Most, especially those with an information theory or cryptography background, use “digital signature” to refer to a digital signature protocol using cryptographic techniques, as is sometimes applied to an 'electronic document'. Many, however, use the terms interchangeably, leading to considerable confusion as cryptographic signature techniques are very different, whatever the term used, than other electronic signatures and have extremely different security properties. Since it is the security properties which are of interest in signatures of all kinds, this is a very significant distinction. Digital signature is properly a subset of electronic signature.

In the European Union, the EU Directive on Electronic Signatures or the EU Electronic Signatures Directive was published in the EC Official Journal, as Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures (OJ No L 13 p.12 19/1/2000).

3.4 Notions of Security

In their foundational paper, Goldwasser, Micali, and Rivest lay out a heirarchy of attack models against digital signatures:

- In a key-only attack, the attacker is only given the public verification key.
- In a known message attack, the attacker is given valid signatures for a variety of messages known by the attacker but not chosen by the attacker.
- In an adaptive chosen message attack, the attacker first learns signatures on arbitrary messages of the attacker's choice.

They also describe a hierarchy of attack results:

- A total break results in the recovery of the signing key.
- A universal forgery attack results in the ability to forge signatures for any message.
- A selective forgery attack results in a signature on a message of the adversary's choice.
- An existential forgery merely results in some valid message/signature pair not already known to the adversary.

The strongest notion of security, therefore, is security against existential forgery under an adaptive chosen message attack.

3.5 Benefits of Digital Signature

Below are some common reasons for applying a digital signature to communications:

- Authentication

Although messages may often include information about the entity sending a message, that information may not be accurate. Digital signatures can be used to authenticate the source of messages. When ownership of a digital signature secret key is bound to a specific user, a valid signature shows that the message was sent by that user. The importance of high confidence in sender authenticity is especially obvious in a financial context. For example, suppose a bank's branch office sends instructions to the central office requesting a change in the

balance of an account. If the central office is not convinced that such a message is truly sent from an authorized source, acting on such a request could be a grave mistake.

• Integrity

In many scenarios, the sender and receiver of a message may have a need for confidence that the message has not been altered during transmission. Although encryption hides the contents of a message, it may be possible to change an encrypted message without understanding it. (Some encryption algorithms, known as nonmalleable ones, prevent this, but others do not.) However, if a message is digitally signed, any change in the message will invalidate the signature. Furthermore, there is no efficient way to modify a message and its signature to produce a new message with a valid signature, because this is still considered to be computationally infeasible by most cryptographic hash functions

3.6 Drawbacks of Digital Signatures

Despite their usefulness, digital signatures alone do not solve the following problems:

• Association of Digital Signatures and Trusted Time Stamping

Digital signature algorithms and protocols do not inherently provide certainty about the date and time at which the underlying document was signed. The signer might have included a time stamp with the signature, or the document itself might have a date mentioned on it. Regardless of the document's contents, a reader cannot be certain the signer did not, for example, backdate the date or time of the signature. Such misuse can be made impracticable by using trusted time stamping in addition to digital signatures.

• Non-Repudiation

In a cryptographic context, the word repudiation refers to any act of disclaiming responsibility for a message. A message's recipient may insist the sender attach a signature in order to make later repudiation more difficult, since the recipient can show the signed message to a third party (e.g., a court) to reinforce a claim as to its signatories' integrity. However, loss of control over a user's private key will mean that all digital signatures using that key, and so ostensibly 'from' that user, are suspect. Nonetheless, a user cannot repudiate a signed message without repudiating their signature key. This is aggravated by the fact there is no trusted time stamp, so new documents (after the key compromise) cannot be separated from old ones, further complicating signature key invalidation. A non-repudiation service requires the existence of a public key infrastructure (PKI) which is complex to establish and operate. The Certificate authorities in a PKI usually maintain a public repository of public keys so the associated private key is certified and signatures cannot be repudiated. Expired certificates are normally removed from the repository. It is a matter for the security policy and the responsibility of the authority to keep old certificates for a period of time if non-repudiation of data service is provided.

• WYSIWYS

Technically speaking, a digital signature applies to a string of bits, whereas humans and applications “believe” that they sign the semantic interpretation of those bits. In order to be semantically interpreted the bit string must be transformed into a form that is meaningful for humans and applications, and this is done through a combination of hardware and software based processes on a computer system. The problem is that

the semantic interpretation of bits can change as a function of the processes used to transform the bits into semantic content. It is relatively easy to change the interpretation of a digital document by implementing changes on the computer system where the document is being processed. From a semantic perspective this creates uncertainty about what exactly has been signed. WYSIWYS (What You See Is What You Sign) means that the semantic interpretation of a signed message can not be changed. In particular this also means that a message can not contain hidden info that the signer is unaware of, and that can be revealed after the signature has been applied. WYSIWYS is a desirable property of digital signatures that is difficult to guarantee because of the increasing complexity of modern computer systems.

3.7 Additional Security Precautions

Putting the private key on a smart card
All public key / private key cryptosystems depend entirely on keeping the private key secret. A private key can be stored on a user's computer, and protected by a local password, but this has two disadvantages:

- the user can only sign documents on that particular computer
- the security of the private key depends entirely on the security of the computer

A more secure alternative is to store the private key on a smart card. Many smart cards are designed to be tamper-resistant (although some designs have been broken, notably by Ross Anderson and his students). In a typical digital signature implementation, the hash calculated from the document is sent to the smart card, whose CPU encrypts the hash using the stored private key of the user, and then returns the encrypted hash. Typically, a user must activate his smart card by entering a personal identification number or PIN code (thus providing two-factor authentication). It can be arranged that the private key never leaves the smart card, although this is not always implemented. If the smart card is stolen, the thief will still need the PIN code to generate a digital signature. This reduces the security of the scheme to that of the PIN system, although it still requires an attacker to possess the card. A mitigating factor is that private keys, if generated and stored on smart cards, are usually regarded as difficult to copy, and are assumed to exist in exactly one copy. Thus, the loss of the smart card may be detected by the owner and the corresponding certificate can be immediately revoked. Private keys that are protected by software only may be easier to copy, and such compromises are far more difficult to detect.

Using Smart card readers With a Separate keyboard

Entering a PIN code to activate the smart card commonly requires a numeric keypad. Some card readers have their own numeric keypad. This is safer than using a card reader integrated into a PC, and then entering the PIN using that computer's keyboard. Readers with a numeric keypad are meant to circumvent the eavesdropping threat where the computer might be running a keystroke logger, potentially compromising the PIN code. Specialized card readers are also less vulnerable to tampering with their software or hardware and are often EAL3 certified.

Other Smart Card Designs

Smart card design is an active field, and there are smart card schemes which are intended to avoid these particular problems, though so far with little security proofs.

Using Digital Signature Only With Trusted Applications

One of the main differences between a digital signature and a written signature is that the user does not "see" what he signs. The application presents a hash code to be encrypted by the digital signing algorithm using the private key. An attacker who gains control of the user's PC can possibly replace the user application with a substitute, in effect replacing the user's own communications with those of the attacker. This could allow a malicious application to trick a user into signing any document by displaying the user's original on-screen, but presenting the attacker's own documents to the signing application.

To protect against this scenario, an authentication system can be set up between the user's application (word processor, email client, etc.) and the signing application. The general idea is to provide some means for both the user app and signing app to verify each other's integrity. For example, the signing application may require all requests to come from digitally-signed binaries.

3.8 Some Digital Signature Algorithms

- Full Domain Hash, RSA-PSS etc., based on RSA
- DSA
- ECDSA
- ElGamal signature scheme
- Undeniable signature
- SHA (typically SHA-1) with RSA
- Rabin signature algorithm

- Pointcheval-Stern signature algorithm
- Schnorr signature
- Aggregate signature-a signature scheme that supports aggregation: Given n signatures on n messages from n users, it is possible to aggregate all these signatures into a single signature whose size is constant in the number of users. This single signature will convince the verifier that the n users did indeed sign the n original messages.

3.9 The Current State of Use-Legal and Practical

Digital signature schemes all have several prior requirements without which no such signature can mean anything, whatever the cryptographic theory or legal provision.

- First, quality algorithms. Some public-key algorithms are known to be insecure, practicable attacks against them having been discovered.
- Second, quality implementations. An implementation of a good algorithm (or protocol) with mistake(s) will not work.
- Third, the private key must remain actually secret; if it becomes known to any other party, that party can produce perfect digital signatures of anything whatsoever.
- Fourth, distribution of public keys must be done in such a way that the public key claimed to belong to, say, Bob actually belongs to Bob, and vice versa. This is commonly done using a public key infrastructure and the public key user association is attested by the operator of the PKI (called a certificate authority). For 'open' PKIs in which anyone can request such an attestation (universally embodied in a cryptographically protected identity certificate), the possibility of mistaken attestation is non trivial. Commercial PKI operators have suffered several publicly known problems. Such mistakes could lead to falsely signed, and thus wrongly attributed, documents. 'closed' PKI systems are more expensive, but less easily subverted in this way.
- Fifth, users (and their software) must carry out the signature protocol properly.

Only if all of these conditions are met will a digital signature actually be any evidence of who sent the message, and therefore of their assent to its contents. Legal enactment cannot change this reality of the existing engineering possibilities, though some such have not reflected this actuality.

Legislatures, being importuned by businesses expecting to profit from operating a PKI, or by the technological avant-garde advocating new solutions to old problems, have enacted statutes and/or regulations in many jurisdictions authorizing, endorsing, encouraging, or permitting

digital signatures and providing for (or limiting) their legal effect. The first and California. Other countries have also passed statutes or issued regulations in this area as well and the UN has had an active model law project for some time. These enactments (or proposed enactments) vary from place to place, have typically embodied expectations at variance (optimistically or pessimistically) with the state of the underlying cryptographic engineering, and have had the net effect of confusing potential users and specifiers, nearly all of whom are not cryptographically knowledgeable. Adoption of technical standards for digital signatures have lagged behind much of the legislation, delaying a more or less unified engineering position on interoperability, algorithm choice, key lengths, and so on what the engineering is attempting to provide.

Using Separate key Pairs for Signing and Encryption

In several countries, a digital signature has a status somewhat like that of a traditional pen and paper signature. Generally, these provisions mean that what is digitally signed legally binds the signer of the document to the terms therein. For that reason, it is often thought best to use separate key pairs for encrypting and signing. Using the encryption key pair, a person can engage in an encrypted conversation (e.g., regarding a real estate transaction), but the encryption does not legally sign every message he sends. Only when both parties come to an agreement do they sign a contract with their signing keys, and only then are they legally bound by the terms of a specific document. After signing, the document can be sent over the encrypted link.

4.0 CONCLUSION

Electronic signature and its offshoot, digital signature has helped tremendously in tracking and minimizing illegal access to databases and accounts. It has also brought about ease and speed in transactions. Though there are attending challenges, but this is far below the advantages.

5.0 SUMMARY

- A digital signature or digital signature scheme is a type of asymmetric cryptography used to simulate the security properties of a handwritten signature on paper.
- In the famous paper “New Directions in Cryptography”, Whitfield Diffie and Martin Hellman first described the notion of a digital signature scheme, although they only conjectured that such schemes existed
- The term electronic signature has several meanings. Among the more expansive is that given by US law, influenced by ABA

committee white papers and the uniform law promulgated by the National Conference of Commissioners on Uniform State Laws (NCCUSL). Under the Uniform Electronic Transactions Act or "UETA" released by NCCUSL in 1999

- In their foundational paper, Goldwasser, Micali, and Rivest lay out a heirarchy of attack models against digital signatures
- Although messages may often include information about the entity sending a message, that information may not be accurate. Digital signatures can be used to authenticate the source of messages.
- Digital signature algorithms and protocols do not inherently provide certainty about the date and time at which the underlying document was signed
- All public key / private key cryptosystems depend entirely on keeping the private key secret. A private key can be stored on a user's computer, and protected by a local password, but this has two disadvantages:
- Digital signature schemes all have several prior requirements without which no such signature can mean anything, whatever the cryptographic theory or legal provision

6.0 TUTOR-MARKED ASSIGNMENT

1. Briefly describe the hierarchy of attack result in digital signature security.
2. List 5 digital signature algorithms.

7.0 REFERENCES/FURTHER READINGS

US ESIGN Act of 2000

The University of Virginia State of WI

National Archives of Australia

"New Directions in Cryptography", IEEE Transactions on Information Theory, IT-22(6):644-654, Nov. 1976.

"Signature Schemes and Applications to Cryptographic Protocol Design", Anna Lysyanskaya, PhD thesis, MIT, 2002.

"A Method For Obtaining Digital Signatures and Public-Key Cryptosystems," Communications of the ACM, 21(2): 120-126, Feb. 1978.

"Constructing Digital Signatures From a One-Way Function.", Leslie Lamport, Technical Report CSL-98, SRI International, Oct. 1979.

"A Certified Digital Signature", Ralph Merkle, In Gilles Brassard, ed., Advances in Cryptology -- CRYPTO '89, vol. 435 of Lecture Notes in Computer Science, pp. 218-238, Spring Verlag, 1990.

"Digitalized signatures as intractable as factorization." Michael O. Rabin, Technical Report MIT/LCS/TR-212, MIT Laboratory for Computer Science, Jan. 1979

"A *Digital Signature Scheme Secure Against Adaptive Chosen-Message Attacks."*, *Shafi Goldwasser, Silvio Micali, and Ronald Rivest.* SIAM Journal on Computing, 17(2):281-308, Apr. 1988.

"Modern Cryptography: Theory & Practice", Wenbo Mao, Prentice Hall Professional Technical Reference, New Jersey, 2004, pg. 308. ISBN 0-13-066943-1

A. Jøsang, D. Povey and A. Ho. "What You See is Not Always What You *Sign"*. *Proceedings of the Australian Unix User Group* Symposium (AUUG2002), Melbourne, September, 2002. PDF

UNIT 2 BIOMETRIC IDENTIFICATION

CONTENTS

1.0 INTRODUCTION

Envision a day when the door to a secured office building can be opened by using an automated system for identification based on a person's physical presence, even though that person left his or her ID or access card on the kitchen counter at home. Imagine ticket-less airline travel, whereby a person can enter the aircraft based on a positive identification verified biometrically at the gateway. Picture getting into a car, starting the engine by flipping down the driver's visor, and glancing into the mirror and driving away, secure in the knowledge that only authorized individuals can make the vehicle operate.

The day when these actions are routine is rapidly approaching. Actually, implementation of fast, accurate, reliable, and user-acceptable biometric identification systems is already underway. Societal behavior patterns result in ever-increasing requirements for automated positive identification systems, and these are growing even more rapidly. The potential applications for these systems are limited only by a person's imagination. Performance claims cover the full spectrum from realistic to incredible. System implementation problems with these new technologies have been predictably high. User acceptance obstacles are on the rise. Security practitioners contemplating use of these systems are faced with overwhelming amounts of often contradictory information provided by manufacturers and dealers.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- trace the development of biometric controls
- identify the applications of biometrics in security management
- identify the characteristics associated with biometrics
- answer the question of problem in deploying biometrics in controls
- explain and differentiate the characteristics of the different types of biometrics.

3.0 MAIN CONTENT

3.1 Biometric Development

Once it became apparent that truly positive identification could only be based on the physical attributes of the person, two questions had to be answered. First, what part of the body could be used? Second, how could identification be accomplished with sufficient accuracy, reliability, and speed so as to be viable in field performance? However, had the pressures demanding automated personal identification not been rising rapidly at the highest levels (making necessary resources and funds available), this research would not have occurred.

At the time, the only measurable characteristic associated with human body that was universally accepted as a positive identifier was the fingerprint. Contact data collected using special inks, dusting powders, and tape, for example, are matched by specially trained experts. Uniquely positioned whorls, ridge endings, and bifurcations were located and compared against templates. A sensor capable of reading a print made by a finger pressed against a piece of glass was required. Matching the collected print against a stored template is a classic computer task. Fortuitously, at the time these identification questions were being asked, computer processing capabilities and speed were increasing rapidly, while size and cost were falling. Had this not been the case, even the initial development of biometric systems would not have taken place. It has taken an additional 25 years of computer and biometric advancement, and cost reduction, for biometrics to achieve widespread acceptability and field proliferation.

Predictably, the early fingerprint-identifying verification systems were not successful in the marketplace, but not because they could not do what they were designed to do. They did. Key problems were the slow decision speed and the lack of ability to detect counterfeit fingerprints. Throughput of two to three persons per minute results in waiting lines, personal frustration, and lost productive time. Failure to detect

counterfeit input (i.e., rubber fingers, photo images) can result in false acceptance of impostors.

Continued comprehensive research and development and advancements in sensing and data processing technologies enabled production of systems acceptable in field use. Even these systems were not without problems, however. Some systems required high levels of maintenance and adjustment for reliable performance. Some required lengthy enrollment procedures. Some required data templates of many thousands of bytes, requiring large amounts of expensive storage media and slowing processing time. Throughput was still relatively slow (though acceptable). Accuracy rates (i.e., false accept and mostly false reject) were higher than would be acceptable today. However, automated biometric identifying verification systems were now performing needed functions in the field.

The value of fast, accurate, and reliable biometric identity verification was rapidly recognized, even if it was not yet fully available. Soon, the number of organized biometric research and development efforts exceeded 20. Many were fingerprint spinoffs: thumb print; full finger print; finger pattern (i.e., creases on the underside of the finger); and palm print. Hand topography (i.e., the side-view elevations of the parts of the hand placed against a flat surface) proved not sufficiently unique for accurate verification, but combined with a top view of the hand (i.e., hand geometry) it became one of the most successful systems in the field. Two-finger geometry is a recently marketed variation.

Other technologies that have achieved at least some degree of market acceptance include voice patterns, retina scan (i.e., the blood-vessel pattern inside the eyeball), signature dynamics (i.e., the speed, direction, and pressure of pen strokes), and iris recognition (i.e., the pattern of features in the colored portion of the eye around the pupil). Others that have reached the market, but have not remained, include keystroke dynamics (i.e., the measurable pattern of speed and time in typing words) and signature recognition (i.e., matching). Other physical characteristics that have been and are currently being investigated as potential biometric identifiers include finger length (though not sufficiently unique), wrist veins (underside), hand veins (back of the hand), knuckle creases (when grasping a bar), fingertip structure (blood vessel pattern under the skin), finger sections (between first and second joint), ear shape, and lip shape. One organization has been spending significant amounts investigating biometric identification based on body odor.

Another biometric identifying verification area receiving significant attention (and funding) is facial recognition. This partially results from

the ease of acquiring facial images with standard video technology and from the perceived high payoff to be enjoyed by a successful facial recognition system. Facial thermograph (i.e., heat patterns of the facial tissue) is an expensive variation because of high camera cost.

The history of the development of biometric identifying verification systems is far from complete. Entrepreneurs continue to see rich rewards for faster, more accurate, and reliable technology, and advanced development will continue. However, advancements are expected to be improvements or variations of current technologies. These will be associated with the hands, eyes, and face for the "what we are" systems and the voice and signature for the "what we do" systems.

3.2 Characteristics of Biometric Systems

These are the important factors necessary for any effective biometric system: accuracy, speed and throughput rate, acceptability to users, uniqueness of the biometric organ and action, resistance to counterfeiting, reliability, data storage requirements, enrollment time, intrusiveness of data collection, and subject and system contact requirements.

Accuracy

Accuracy is the most critical characteristic of a biometric identifying verification system. If the system cannot accurately separate authentic persons from impostors, it should not even be termed a biometric identification system.

False Reject Rate: The rate, generally stated as a percentage, at which authentic, enrolled persons are rejected as unidentified or unverified persons by a biometric system is termed the false reject rate. False rejection is sometimes called a Type I error. In access control, if the requirement is to keep the "bad guys" out, false rejection is considered the least important error. However, in other biometric applications, it may be the most important error. When used by a bank or retail store to authenticate customer identity and account balance, false rejection means that the transaction or sale (and associated profit) is lost, and the customer becomes upset. Most bankers and retailers are willing to allow a few false accepts as long as there are no false rejects.

False rejections also have a negative effect on throughput, frustrations, and unimpeded operations, because they cause unnecessary delays in personnel movements. An associated problem that is sometimes incorrectly attributed to false rejection is failure to acquire. Failure to acquire occurs when the biometric sensor is not presented with sufficient usable data to make an authentic or impostor decision. Examples include

smudged prints on a fingerprint system, improper hand positioning on a hand geometry system, improper alignment on a retina or iris system, or mumbling on a voice system. Subjects cause failure to acquire problems, either accidentally or on purpose.

False Accept Rate: The rate generally stated as a percentage, at which unenrolled or impostor persons are accepted as authentic, enrolled persons by a biometric system is termed the false accept rate. False acceptance is sometimes called a Type II error. This is usually considered to be the most important error for a biometric access control system.

Crossover Error Rate (CER): This is also called the equal error rate and is the point, generally stated as a percentage, at which the false rejection rate and the false acceptance rate are equal. This has become the most important measure of biometric system accuracy.

All biometric systems have sensitivity adjustment capability. If false acceptance is not desired, the system can be set to require (nearly) perfect matches of enrollment data and input data. If tested in this configuration, the system can truthfully be stated to achieve a (near) zero false accept rate. If false rejection is not desired, this system can be readjusted to accept input data that only approximate a match with enrollment data. If tested in this configuration, the system can be truthfully stated to achieve a (near) zero false rejection rate. However, the reality is that biometric systems can operate on only one sensitivity setting at a time.

The reality is also that when system sensitivity is set to minimize false acceptance, closely matching data will be spurned, and the false rejection rate will go up significantly. Conversely, when system sensitivity is set to minimize false rejects, the false acceptance rate will go up notably. Thus, the published (i.e., truthful) data tell only part of the story. Actual system accuracy in field operations may even be less than acceptable. This is the situation that created the need for a single measure of biometric system accuracy.

The crossover error rate (CER) provides a single measurement that is fair and impartial in comparing the performance of the various systems. In general, the sensitivity setting that produces the equal error will be close to the setting that will be optimal for field operation of the system. A biometric system that delivers a CER of 2% will be more accurate than a system with a CER of 5%.

Speed and Throughput Rate

The speed and throughput rate are the most important biometric system characteristics. Speed is often related to the data processing capability of the system and is stated as how fast they accept or reject decision is annunciated. In actuality, it relates to the entire authentication procedure: stepping up to the system; inputting the card or PIN (if a verification system); input of the physical data by inserting a hand or finger, aligning an eye, speaking access words, or signing a name; processing and matching of data files; annunciation of the accept or reject decision; and, if a portal system, movement through and closing the door.

Generally accepted standards include a system speed of 5 seconds from startup through decision annunciation. Another standard is a portal throughput rate of 6 to 10/minute, which equates to 6 to 10 seconds/person through the door. Only in recent years have biometric systems become capable of meeting these speed standards, and, even today, some marketed systems do not maintain this rapidity. Slow speed and the resultant waiting lines and movement delays have frequently caused the removal of biometric systems and even the failure of biometric companies.

Acceptability to Users

System acceptability to the people who must use it has been a little noticed but increasingly important factor in biometric identification operations. Initially, when there were few systems, most were of high security and the few users had a high incentive to use the systems; user acceptance was of little interest. In addition, little user threat was seen in fingerprint and hand systems.

Biometric system acceptance occurs when those who must use the system — organizational managers and any union present — all agree that there are assets that need protection, the biometric system effectively controls access to these assets, system usage is not hazardous to the health of the users, system usage does not inordinately impede personnel movement and cause production delays, and the system does not enable management to collect personal or health information about the users. Any of the parties can effect system success or removal. Uncooperative users will overtly or covertly compromise, damage, or sabotage system equipment. The cost of union inclusion of the biometric system in their contracts may become too costly. Moreover, management has the final decision on whether the biometric system benefits outweigh its liabilities.

Uniqueness of Biometric Organ and Action

Because the purpose of biometric systems is positive identification of personnel, some organizations (e.g., elements of the government) are specifying systems based only on a unique (i.e., no duplicate in the world) physical characteristic. The rationale is that when the base is a unique characteristic, a file match is a positive identification rather than a statement of high probability that this is the right person. Only three physical characteristics or human organs used for biometric identification are unique: the fingerprint, the retina of the eye (i.e., the blood-vessel pattern inside the back of the eyeball), and the iris of the eye (i.e., random pattern of features in the colored portion of the eye surrounding the pupil). These features include freckles, rings, pits, striations, vasculature, coronas, and crypts.

Resistance to Counterfeiting

The ability to detect or reject counterfeit input data is vital to a biometric access control system meeting high security requirements. These include use of rubber, plastic, or even hands or fingers of the deceased in hand or fingerprint systems, and mimicked or recorded input to voice systems. Entertainment media, such as the James Bond or Terminator films, have frequently shown security system failures when the heads or eyes of deceased (i.e., authentic) persons were used to gain access to protected assets or information. Because most of the early biometric identifying verification systems were designed for high security access control applications, failure to detect or reject counterfeit input data was the reason for several system or organization failures. Resistance to counterfeit data remains a criterion of high-quality, high-accuracy systems. However, the proliferation of biometric systems into other non-high-security type applications means that lack of resistance to counterfeiting is not likely to cause the failure of a system in the future.

Reliability

It is vital that biometric identifying verification systems remain in continuous, accurate operation. The system must allow authorized persons access while precluding others, without breakdown or deterioration in performance accuracy or speed. In addition, these performance standards must be sustained without high levels of maintenance or frequent diagnostics and system adjustments.

Data Storage Requirements

Data storage requirements are a far less significant issue today than in the earlier biometric systems when storage media were very expensive. Nevertheless, the size of biometric data files remains a factor of interest. Even with current ultra-high-speed processors, large data files take

longer to process than small files, especially in systems that perform full identification, matching the input file against every file in the data base. Biometric file size varies between 9 and 10,000 bytes, with most falling in the 256- to 1,000-byte range.

Enrollment Time

Enrollment time is also a less significant factor today. Early biometric systems sometimes had enrollment procedures requiring many repetitions and several minutes to complete. A system requiring a 5-minute enrollment instead of 2 minutes causes 50 hours of expensive nonproductive time if 1,000 users must be enrolled. Moreover, when line waiting time is considered, the cost increases several times. The accepted standard for enrollment time is 2 minutes per person. Most of the systems in the marketplace today meet this standard.

Intrusiveness of Data Collection

Originally, this factor developed because of user concerns regarding collection of biometric data from inside the body, specifically, the retina inside the eyeball. Early systems illuminated the retina with a red light beam. However, this coincided with increasing public awareness of lasers, sometimes demonstrated as red light beams cutting steel. There has never been an allegation of user injury from retina scanning, but user sensitivity expanded from resistance to red lights intruding inside the body to include any intrusion inside the body. This user sensitivity has now increased to concerns about intrusions into perceived personal space.

Subject and System Contact Requirements

This factor could possibly be considered as a next step or continuation of intrusiveness. Indications are that biometric system users are becoming increasingly sensitive to being required to make firm physical contact with surfaces where up to hundreds of other unknown (to them) persons are required to make contact for biometric data collection.

These concerns include voice systems that require holding and speaking into a handset close to the lips.

There seems to be some user feeling that: “if I choose to do something, it is OK, but if an organization, or society, requires me to do the same thing, it is wrong.” Whether or not this makes sense, it is an attitude spreading through society that is having an impact on the use of biometric systems. Systems using video camera data acquisition do not fall into this category.

3.3 Biometric Problems

A variety of problems in the field utilization of biometric systems over the past 25 years have been identified. Some have been overcome and are seldom seen today; others still occur. These problems include performance, hardware and software robustness, maintenance requirements, susceptibility to sabotage, perceived health maladies because of usage, private information being made available to management, and skill and cooperation required to use the system.

Performance

Field performance of biometric identifying verification systems is often different than that experienced in manufacturers' or laboratory tests. There are two ways to avoid being stuck with a system that fails to deliver promised performance. First, limit consideration to technologies and systems that have been tested by an independent, unbiased testing organization. Sandia National Laboratories, located in Albuquerque, New Mexico, has done biometric system testing for the Department of Energy for many years, and some of their reports are available. Second, any system manufacturer or sales representative should be able to provide a list of organizations currently using their system. They should be able to point out those users whose application is similar to that currently contemplated (unless the planned operation is a new and unique application). Detailed discussions, and perhaps a site visit, with current users with similar application requirements should answer most questions and prevent many surprises.

Hardware and Software Robustness

Some systems and technologies that are very effective with small- to medium-sized user data bases have a performance that is less than acceptable with large data bases. Problems that occur include system slowdown and accuracy degradation. Some biometric system users have had to discard their systems and start over because their organizations became more successful grew faster than anticipated, and the old system could not handle the growth. If they hope to "grow" their original system with the organization, system managers should at least double the most optimistic growth estimate and plan for a system capable of handling that load.

Another consideration is hardware capability to withstand extended usage under the conditions expected. An example is the early signature dynamics systems, which performed adequately during testing and early fielding periods. However, the pen and stylus sensors used to detect stroke direction, speed, and pressures were very tiny and sensitive. After months or a year of normal public use, the system performance had

deteriorated to the point that the systems were no longer effective identifiers.

Maintenance Requirements

Some sensors and systems have required very high levels of preventive maintenance or diagnostics and adjustment to continue effective operations. Under certain operating and user conditions (e.g., dusty areas or with frequent users of hand lotions or creams), some fingerprint sensors needed cleaning as frequently as every day to prevent deterioration of accuracy. Other systems demanded weekly or monthly connection of diagnostic equipment, evaluation of performance parameters, and careful adjustment to retain productive performance. These human interventions not only disrupt the normal security process, but significantly increase operational costs.

Susceptibility to Sabotage

Systems with data acquisition sensors on pedestals protruding far out from walls or with many moving parts are often susceptible to sabotage or disabling damage. Spinning floor polisher handles or hammers projecting out of pockets can unobtrusively or accidentally affect sensors. These incidents have most frequently occurred when there was widespread user or union resistance to the biometric system.

Perceived Health Maladies Due to Usage

As new systems and technologies were developed and public sensitivity to new viruses and diseases such as AIDS, Ebola, and E. coli increased by orders of magnitude, acceptability became a more important issue. Perceptions of possible organ damage and potential spread of disease from biometric system usage ultimately had such a devastating affect on sales of one system that it had to be totally redesigned. Though thousands of the original units had been successfully fielded, whether the newly packaged technology regains popularity or even survives remains to be seen. All of this occurred without even one documented allegation of a single user becoming sick or injured as a result of system utilization.

Many of the highly contagious diseases recently publicized can be spread by simple contact with a contaminated surface. As biometric systems achieve wider market penetration in many applications, user numbers are growing logarithmically. There are developing indications that users are becoming increasingly sensitive about systems and technologies that require firm physical contact for acquisition of the biometric data.

Private Information Made Available to Management

Certain health events can cause changes in the blood vessel pattern (i.e., retina) inside the eyeball. These include diabetes and strokes. Allegations have been made that the retina-based biometric system enables management to improperly obtain health information that may be used to the detriment of system users. The scenario begins with the system failing to identify a routine user. The user is easily authenticated and re-enrolled. As a result, management will allegedly note the re-enrollment report and conclude that this user had a minor health incident (minor because the user is present the next working day). In anticipation that this employee's next health event could cause major medical cost, management might find (or create) a reason for termination. Despite the fact that there is no recorded case of actual occurrence of this alleged scenario, this folklore continues to be heard within the biometric industry.

Skill and Cooperation Required Using the System

The performance of some biometric systems is greatly dependent on the skill or careful cooperation of the subject in using the system. Though there is an element of this factor required for data acquisition positioning for all biometric systems, it is generally attributed to the "what we do" type of systems.

3.4 Benefits of Biometric Identification As Compared With Card Systems

Biometric identifying verification systems control people. If the person with the correct hand, eye, face, signature, or voice is not present, the identification and verification cannot take place and the desired action (i.e., portal passage, data, or resource access) does not occur.

As has been demonstrated many times, adversaries and criminals obtain and successfully use access cards, even those that require the addition of a PIN. This is because these systems control only pieces of plastic (and sometimes information), rather than people. Real asset and resource protection can only be accomplished by people, not cards and information, because unauthorized persons can (and do) obtain the cards and information.

Further, life-cycle costs are significantly reduced because no card or PIN administration system or personnel are required. The authorized person does not lose physical characteristics (i.e., hands, face, eyes, signature, or voice), but cards and PINs are continuously lost, stolen, or forgotten. This is why card access systems require systems and people to

administer, control, record, and issue (new) cards and PINs. Moreover, the cards are an expensive and recurring cost.

Card System Error Rates

The false accept rate is 100% when the access card is in the wrong hands, lost, or stolen. It is a false reject when the right card is swiped incorrectly or just does not activate the system. (Think about the number of times to retry hotel room access cards to get the door to unlock.) Actually, it is also a false reject when a card is forgotten and that person cannot get through the door.

3.7 Different Types of Biometric systems and Their Characteristics

This sub unit describes the different types of biometric systems: fingerprint systems, hand geometry systems, voice pattern systems, retina pattern systems, iris pattern systems, and signature dynamics systems. For each system these characteristics are described: the enrollment procedure and time, the template or file size, the user action required, the system response time, any anticounterfeit method, accuracy, field history, problems experienced, and unique system aspects.

Fingerprint Systems
The information in this section is a compilation of information about several biometric identifying verification systems whose technology is based on the fingerprint.

Data Acquisition: Fingerprint data is acquired when subjects firmly press their fingers against a glass or polycarbonate plate. The fingerprint image is not stored. Information on the relative location of the ridges, whorls, lines, bifurcations, and intersections is stored as an enrolled user data base file and later compared with user input data.

User Actions Required: Nearly all fingerprint-based biometrics are verification systems. The user states identification by entering a PIN through a keypad or by using a card reader, then places a finger on the reader plate.

Accuracy: Some fingerprint systems can be adjusted to achieve a false accept rate of 0.0%. Sandia National Laboratories tests of a top-rated fingerprint system in 1991 and 1993 produced a three-try false reject rate of 9.4% and a crossover error rate of 5%.

Problems Experienced: System operators with large user populations are often required to clean sensor plates frequently to remove built-up skin oil and dirt that adversely affect system accuracy.

Hand Geometry System

Hand geometry data, the three-dimensional record of the length, width, and height of the hand and fingers is acquired by simultaneous vertical and horizontal camera images.

User Actions Required: The hand geometry system operates only as an identification verifier. The user states identification by entering a PIN on a keypad or by using a card reader. When the "place hand" message appears on the unit display, the user places the hand flat on the platen against the pegs. When all four lights confirm correct hand position the data are acquired and a "remove hand" message appears.

Accuracy: Sandia National Laboratories tests have produced a one-try false accept rate less than 0.1%, a three-try false reject rate less than 0.1%, and crossover error rates of 0.2 and 2.2% (i.e., two tests).

Problems Experienced: Some of the field applications did not perform up to the accuracy results of the initial Sandia test. There have been indications that verification accuracy achieved when user data bases are in the hundreds deteriorates when the data base grows into the thousands.

Voice Pattern Systems
Up to seven parameters of nasal tones, larynx and throat vibrations, and air pressure from the voice are captured by audio and other sensors.

User Actions Required: Currently, voice systems operate only as identification verifiers. The user states identification by entering the PIN on the telephone-type keypad. As cued through the handset (i.e., recorded voice stating "please say your access phrase"), the user speaks into the handset sensors.

Accuracy: Sandia National Laboratories has reported crossover errors over 10% for two systems they have tested. Other voice tests are being planned.

Problems Experienced: Background noise can affect the accuracy of voice systems. Access systems are located at entrances, hallways, and doorways, which tend to be busy, high-traffic, and high-noise-level sites.

Retina Pattern System

The system records elements of the blood-vessel pattern of the retina on the inside rear portion of the eyeball by using a camera to acquire the image.

User Actions Required: If verifying, the user enters the PIN on the keypad. The system automatically acquires data when an eye is positioned in front of the aperture and centered on the pulsing green dot. Acceptance or nonacceptance is indicated in the LCD display.

Accuracy: Sandia National Laboratories test of the previous retina model produced no false accepts and a crossover error rate of 1.5%. The new model, System 2001, is expected to perform similarly.

Field History: Hundreds of the original binocular-type units were fielded before those models were discontinued. They were used for access control and identification in colleges, laboratories, government facilities, and jails. The new model, System 2001, is now on sale.

Problems Experienced: Because persons perspiring or having watery eyes could leave moisture on the eyecups of the previous models, some users were concerned about acquiring a disease through the transfer of body fluids. Because the previous models used a red light beam to acquire pattern data, some users were concerned about possible eye damage from the "laser." No allegations were made that any user actually became injured or diseased through the use of these systems. Because some physical conditions such as diabetes and heart attacks can cause changes in the retinal pattern, which can be detected by the system, some users were concerned that management would gain unauthorized medical information that could be used to their detriment. No cases of detrimental employee personnel actions resulting from retina system information have been reported.

Iris Pattern System

The iris (i.e., the colored portion of the eye surrounding the pupil) has rich and unique patterns of striations, pits, freckles, rifts, fibers, filaments, rings, coronas, furrows, and vasculature. The images are acquired by a standard 1/3 inch CCD video camera capturing 30 images per second, similar to a camcorder..

User Actions Required: The IriScan system can operate as a verifier, but is normally used in full identification mode because it performs this function faster than most systems verify. The user pushes the start

button, tilts the optical unit if necessary to adjust for height, and looks at the LCD feedback image of his or her eye, centering and focusing the image. If the system is used as a verifier, a keypad or cardreader is interconnected.

Accuracy: Sandia National Laboratories' test of a preproduction model had no false accepts, low false rejects, and the system "performed extremely well." Sandia has a production system currently in testing. British Telecommunications recently tested the system in various modes and will publish a report in its engineering journal. They report 100% correct performance on over 250,000 IrisCode comparisons. "Iris recognition is a reliable and robust biometric. Every eye presented was enrolled. There were no False Accepts, and every enrolled eye was successfully recognized." Other tests have reported a crossover error rate of less than 0.5%.

Problems Experienced: Because this is a camera-based system, the optical unit must be positioned such that the sun does not shine directly into the aperture.

Signature Dynamics Systems

The signature pen-stroke speed, direction, and pressure are recorded by small sensors in the pen, stylus, or writing tablet.

User Actions Required: The user states identification through PIN entry on a keypad or cardreader. The signature is then written by using the instrument or tablet provided. Some systems permit the use of a stylus without paper if a copy of the signature is not required for a record.

Accuracy: Data collection is underway at pilot projects and beta test sites. Current signature dynamics biometric systems have not yet been tested by an independent agency.

Field History: Approximately 100 units are being used in about a dozen systems operated by organizations in the medical, pharmaceutical, banking, manufacturing, and government fields.

Problems Experienced: Signature dynamics systems which previously performed well during laboratory and controlled tests, did not stand up to rigorous operational field use. Initially acceptable accuracy and reliability rates began to deteriorate after months of system field use. Although definitive failure information is not available, it is believed that the tiny, super-accurate sensors necessary to measure the minute changes in pen speed, pressure, and direction did not withstand the

rough handling of the public. It is too early to tell whether the current generation of signature systems has overcome these shortcomings.

4.0 CONCLUSION

The era of fast, accurate, cost-effective biometric identification systems has arrived. Societal activities increasingly threaten individual's and organization's assets, information, and, sometimes, even their existence. Instant, positive personal identification is a critically important step in controlling access to and protecting society's resources. Effective tools are now available.

There are more than a dozen companies manufacturing and selling significant numbers of biometric identification systems today. Even more organizations are conducting biometric research and development and hoping to break into the market or already selling small numbers of units. Not all biometric systems and technologies are equally effective in general, nor specifically in meeting all application requirements. Security managers are advised to be cautious and thorough in researching candidate biometric systems before making a selection. Independent test results and the reports of current users with similar applications are recommended. On-site tests are desirable. Those who are diligent and meticulous in their selection and installation of a biometric identification system will realize major increases in asset protection levels.

5.0 SUMMARY

- Envision a day when the door to a secured office building can be opened by using an automated system for identification based on a person's physical presence, even though that person left his or her ID or access card on the kitchen counter at home.
- Once it became apparent that truly positive identification could only be based on the physical attributes of the person, two questions had to be answered. First, what part of the body could be used? Second, how could identification be accomplished with sufficient accuracy, reliability, and speed so as to be viable in field performance?
- These are the important factors necessary for any effective biometric system: accuracy, speed and throughput rate, acceptability to users, uniqueness of the biometric organ and action, resistance to counterfeiting, reliability, data storage requirements, enrollment time, intrusiveness of data collection, and subject and system contact requirements.
- A variety of problems in the field utilization of biometric systems over the past 25 years have been identified. Some have been overcome and are seldom seen today; others still occur.

- Biometric identifying verification systems control people. If the person with the correct hand, eye, face, signature, or voice is not present, the identification and verification cannot take place and the desired action (i.e., portal passage, data, or resource access) does not occur.
- There are different types of biometric systems: fingerprint systems, hand geometry systems, voice pattern systems, retina pattern systems, iris pattern systems, and signature dynamics systems.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention 10 characteristics of a biometric system.
2. Briefly discuss Maintenance Requirements as a problem associated with the use of biometrics in security.

7.0 REFERENCES/FURTHER READINGS

Krause, M. and Tipton, H.F. Handbook of Information Security *Management.*

UNIT 3 FRAUD PREVENTION

CONTENTS

1.0 INTRODUCTION

This unit examines the risks associated with conducting commercial transactions through the use of electronic technologies and how best to prevent funds from being illegally misappropriated in the world of electronic commerce. It will not, however, deal with the many other problems which arise out of on-line commerce such as those relating to misleading and deceptive practices which create particular concerns for consumer protection agencies. Instead, it focuses on the ways of preventing illegality which arises out of the use of electronic payment systems. The extent to which individuals are using the Internet for business transactions is increasing enormously. The market is potentially great with approximately 171.25 million people estimated to be using the Internet worldwide in May 1999 (NUA Internet Surveys 1999a). In terms of commercial usage, Forrester Research has estimated that global business-to-business on-line commerce could amount to US $327 billion by the year 2002. Although the rate of increase in on-line commerce has slowed somewhat recently, it will still represent an important part of economic life in the years to come.

Surveys of Internet usage have shown that most transactions which take place on-line involve small-value purchases such as books, CDs, wine, computers, and information technology products. The potential exists, however, for anything to be purchased electronically and we have recently seen the establishment of a number of on-line auction houses which deal in much higher-value goods A large proportion of Internet users also arrange travel and holidays electronically. Jupiter Communications (1999) has, for example, estimated that the on-line travel market will be worth US$16.6 billion by 2003.
The Internet is, arguably, at the moment an expansive advertising medium in which information as well as goods and services are made available to users throughout the world. Most business transactions take place by purchasers identifying goods and services which they require

by inspecting Internet sites. They are then able to pay for the product chosen using conventional forms of payment, such as money orders or cheques which may be sent by post to the merchant before the product is dispatched or the service provided. Alternatively, purchasers may pay for a product or service by transferring funds electronically. This can be done by disclosing one's bank account (usually a credit card account) details and authorising the merchant to debit the specified account to the value of the sum in question. In this way one could, for example, sitting in Malta, purchase a book from an antiquarian bookshop in London, or subscribe to an adult Internet site emanating from Sydney.

Preventing on-line fraud will involve the use of both conventional as well as novel technological approaches. The essential elements of each are as follows.

2.0 OBJECTIVES

At the end of this unit you should be able to:

- identify the risks involved in e-business
- know the different types of defense strategies against frauds
- differentiate the different forms of strategies
- explain the actions that constitutes a technological control strategy.

3.0 MAIN CONTENT

3.1 Risk Awareness and User Education

One of the most effective strategies used to control crime is education of the public as to the nature of the security risks which they face, and how they may best protect themselves. As most on-line payment systems will require the use of a PIN or password in order for users to gain access to Personal Computers or plastic cards, protection of access codes will be the primary crime prevention strategy available. Users are best placed to protect themselves by taking basic security precautions to ensure that cards are not stolen. This includes not leaving them in public places and ensuring that they are reclaimed after use. Consumers are also advised not to compromise their security by disclosing access codes, keeping them with cards, or writing them on cards. Studies reveal, nonetheless, that between twenty and seventy per cent of people write their PIN on the card or on a piece of paper carried with the card (Sullivan 1987).
In the United Kingdom, one particularly effective plastic card fraud prevention strategy called 'Cardwatch' involved a high profile publicity and education campaign by the Association for Payment Clearing Service including posters, leaflets, and television and radio coverage to raise public awareness of the problem and to encourage card holders to

take more care of their cards. It resulted in a forty-one per cent reduction in card fraud overall between 1991 and 1994, while losses occurring at retail points of sale were reduced by forty-nine per cent during the same period. Losses from cards lost or stolen in the post were reduced by sixty-two per cent between 1991 and 1994 (Webb 1996). Had these fraud prevention initiatives not been introduced, it has been estimated that losses in Britain would have been 350 per cent higher than those recorded in 1996 (see Levi and Handley 1998). When funds transfer systems become fully operational on the Internet, there will also be a need for users to be educated as to ways in which they may protect themselves from fraud and deception. The Internet, itself, may prove to be the most effective mechanism for transmitting such information.

3.2 Institutional Practices

Institutions involved in maintaining the infrastructure of the Internet as well as financial institutions are able to adopt a wide variety of self-help strategies which may reduce the risk of funds transfer fraud on the Internet.

First, and most importantly, is the need for organisations to be confident that the staff they are employing are reliable and trustworthy, as Internet fraud often involves confederates with inside knowledge of an institution's security and computer procedures. Personnel should also be regularly monitored in terms of their risk of behaving fraudulently, particularly long-term employees who have acquired considerable knowledge of an organisation's security procedures. Caution is also needed when organisational disputes develop.

A case heard before the New South Wales District Court on 27 March 1998, for example, concerned an unsuccessful applicant for a position with an Internet Service Provider (ISP). When he was refused the job he took revenge by illegally obtaining access to the company's database of credit card holders and publishing details relating to 1,225 cardholders on the Internet as a demonstration of the security weaknesses of the company. As a result, the business lost more than $A2 million and was forced to close its ISP activities (R. v Stevens unreported decision of the New South Wales District Court, 27 March 1998).

Financial institutions may be able to assist on-line merchants by notifying them of incidents of fraud as soon as they occur in order that they may be able to avoid repeat victimisation or that others may be able to avoid victimisation. In the United Kingdom, for example, a National Hot Card File was created by which details of lost and stolen cards were able to be quickly transmitted to retail outlets. A similar notification

system could be established for on-line merchants. Systems may also need to be created by which on-line merchants are able to obtain immediate authorisation from financial institutions before transactions are accepted. It may even be necessary for all on-line transactions to be authorised before they are accepted.

Frauds in which merchants are involved constitute a large problem for financial institutions as merchants or their employees are ideally placed to permit access to computer networks and to alter transaction details. Financial institutions may need to make use of artificial neural networks in order to isolate fraudulent claiming patterns by merchants and maintain databases of merchants who have engaged in illegal conduct on the Internet in the past. Already, organisations are providing certification services to enable users to identify illegal or unsafe Internet sites. One example is the United States Better Business Bureau which approves safe Web sites for Internet commerce

3.3 Technological Solutions

A wide range of technological solutions have been devised in order to reduce the security risks associated with conducting on-line business.

Hardware Security

In order to provide a safe system for electronic commerce, computer hardware needs to be adequately secured. This extends from computer terminals used in homes, businesses, and public kiosks through servers operated by ISPs, to the hardware maintained by merchants and financial institutions. The extent of the security precautions used will be determined by the risks present. Terminals located in Internet kiosks may need only basic access controls such as through the use of passwords or smartcard tokens, whilst servers maintained by banks might need to be shielded against electro-magnetic radiation (EMR) scanning.

The threat of EMR scanning should not be taken lightly. Although the risk is remote, the possibility exists. In one case in England, for example, a computer eavesdropper scanned electronic transaction information transmitted by a bank. Despite the fact that the information was encrypted, the code was defeated and the individual successfully obtained £350,000 by blackmailing the bank and several customers by threatening to reveal certain information to the Inland Revenue (Nicholson 1989). If payment systems are used which make use of digital signatures and encrypted data transmissions, then the need to protect computer cables from interception would not arise as any data would not travel in clear text. At present, however, a good deal of

sensitive information travels across networks in unencrypted form making it vulnerable to interception and subsequent disclosure. The adequacy of encryption as a security measure depends, of course, upon the strength of the encryption system used and the determination of the attacker.

Terminal Safeguards

Crime prevention needs to be focused on areas of particular weakness in electronic systems and the most obvious target for electronic funds transfer systems is the computer terminal at which transactions are carried out. As is the case with telephone kiosks, ATM and EFTPOS terminals need to be manufactured in such a way as to ensure that access cannot be gained to cables or to electro-magnetic radiation (Tyree 1990). Computer terminals should be located in secure places where users are protected both physically, as well as against shoulder surfing, to obtain PINs. Such safeguards may be easier to enforce where Internet transactions are involved, although some transactions will be carried out in public places such as Internet cafes which will be more difficult to protect.

Card Security

Plastic cards may be used in conjunction with on-line transactions in a variety of ways. Primarily they will be used to store access devices such as cryptographic keys or other user authentication devices. They may also be used to store value in Mondex-type smart card systems. The most sophisticated security features should be built into plastic cards in order to prevent counterfeiting, alteration or unauthorised access to the data which they hold. Newton (1995) describes various crime prevention strategies which have been used to prevent plastic card counterfeiting including the use of security printing, micro-printing, holograms, embossed characters, tamper-evident signature panels, magnetic stripes with improved card validation technologies, and indent printing. Smart cards, of course, are much more difficult to copy than ordinary magnetic stripe cards.

Unfortunately, all of these card security features have been overcome by organised criminals including computer chip circuitry in smart cards. On-line payment systems which do not rely upon plastic cards will, presumably, be much more secure and it may also be possible for these systems to operate in conjunction with biometric user identification systems.

Value Restrictions

As an alternative to target hardening, it has been suggested that the risk of large-scale fraud and money laundering using Internet-based funds transfer systems could be restricted by placing limits on the size of transactions. Mackrell (1996), for example, has suggested that stored value cards should have a modest limit placed on the maximum value that can be stored on them, especially if they are to be used for card-to-card transfers. There could also be a limit on the life of the cards which would restrict their usefulness for hoarding and money laundering. Self-expiring cards have also been developed which automatically deteriorate after a certain period of time. In the case of on-line commerce, electronic restrictions could be placed on the value of transactions in order to avoid the possibility of large scale fraud, although this may be seen as an unwarranted intrusion into freedom of electronic commerce.

Password Protection

Passwords used as a means of restricting access to computer technologies are popular at present and frequently misused and abused. It is possible to guess passwords, particularly if little or no thought has been given to their selection, or to use various forms of social engineering to trick users into revealing their passwords for subsequent improper use.

The use of brute computing force has also been used to break passwords. Password cracking programs are available by which computers are able systematically to search entire dictionaries in search of a password. Even if passwords are encrypted so as to prevent them from direct exposure, encryption keys have been broken through the use of massive computing resources. Denning (1998: 40) reports, for example, that in 1994 a 129 digit RSA key was broken through combining the power of 1,600 computers linked through the Internet globally working for eight months at the rate of one million instructions per second. If additional information or cracks within the system are known, it is possible to break encryption keys even more quickly, which has also been documented. There are various ways of enhancing access security through the use of passwords (see Alexander 1995).

Appropriate education of users is an initial first step in which information is given concerning ways of ensuring that passwords are not disclosed, guessed, or otherwise compromised by the user in question. Systems should be used which change passwords regularly, or which deny access after a specified number of consecutive tries using invalid passwords. Terminals should have automatic shutdown facilities when they have not been used for specified periods, such as five minutes. Single use passwords, where the password changes with every successive login according to an agreed protocol known to the user and

system operator, could also be used. The Secure ID card, for example, generates a new password every sixty seconds which is a function of the time and a secret 64-bit seed that it unique to the card (Denning 1998: 44).

Challenge-response protocols may also be used as a means of carrying out user authentication. The server generates a random number which is sent to the card. In a public key system, the card digitally signs the number and returns it to the server. The server then validates the digital signature. Alternatively, call-back devices may be used. After the user dials into a computer through a modem and gives his or her identity, the system disconnects the user and then telephones the user on a number previously registered with the server. After the user is verified, the transaction can then proceed (see, for example, NetCrusader 1998). Such a system is, however, able to be overcome through the use of call-forwarding arrangements (Denning 1998: 45). Another user authentication system makes use of space geodetic methods to authenticate the physical locations of users, network nodes, and documents. One company, CyberLocator, involves the use of a location signature sensor which uses signals transmitted by satellite to provide a location on earth at any given time. Users are thus able to be located at the time they attempt to gain access to the system, which provides a safeguard against individuals pretending to be legitimate users who are located in a different physical location (Denning 1998: 45).

Cardholder Verification

One of the greatest areas of risk associated with electronic funds transfer systems relates to the manner in which users' identities are verified. Some of the most recent suggestions for improving security in this area include the use of various biometric means of verifying identity such as signature, fingerprint, palm, lip, ear or retina scanning (Sullivan 1987). Masuda (1996) provides an examination of a credit card crime prevention strategy employed since 1993 by Tops Appliance City Inc. in New York called 'Cardwatch'. This involves a computer network in a chain of retail stores in which credit card applications are checked by photographing the applicant digitally, recording the applicant's signature and other identifying information such as driver's licence, telephone and social security numbers, present address and current or last place of employment. This information is then used for future purchases and also when the customer collects merchandise.

Such an approach employs two fundamental checks on identity: something an account holder possesses (the card) and something that an account holder is (photograph etc).

Because information is recorded about the individual, offenders are reluctant to take out accounts fraudulently. Cardwatch resulted in a ninety per cent reduction in credit card fraud losses over a seventeen-month period following introduction of the scheme, with a fifty-seven per cent reduction in per fraud loss.

Value Restrictions

As an alternative to target hardening, it has been suggested that the risk of large-scale fraud and money laundering using electronic funds transfer systems could be restricted by placing limits on the size of transactions. Mackrell (1996), for example, has suggested that stored value cards should have a modest limit placed on the maximum value that can be stored on them, especially if they are to be used for card-to-card transfers. There could also be a limit on the life of the cards which would restrict their usefulness for hoarding and money laundering. In the case of Internet commerce, electronic restrictions could be placed on the value of transactions in order to avoid the possibility of large scale fraud, although this may be seen as an unwarranted intrusion into freedom of electronic commerce

Protections against Card Counterfeiting

Newton (1995) describes various crime prevention strategies which have been used to prevent plastic card counterfeiting. These include the use of security printing; microprinting; holograms; embossed characters; tamper-evident signature panels; magnetic stripes with improved card validation technologies and indent printing. Smart cards, of course, are much more difficult to copy than ordinary magnetic stripe cards.

Unfortunately, all of these card authentication devices have been overcome by organized criminals except for computer chip circuitry in smart cards, which has yet to be fully counterfeited successfully. Internet payment systems which do not rely upon plastic cards will, presumably, be much more secure and it may also be possible for these to operate in conjunction with biometric user identification systems.

Biometrics

One way in which problems of password and token security may be overcome, is for users to identify themselves biometrically. Already there are a wide variety of such systems being used which make use of an individual's unique physical properties. Common biometric identifiers today include fingerprints, voice patterns, typing patterns, retinal images, facial or hand geometry, and even the identification of a person's subcutaneous vein structures or body odours (Johnson 1996).

The body odour system called 'Scentinel' was developed by the British firm Bloodhouse Sensors and requires that you pass your hand under a sensor which records your unique smell and compares it with one registered in the database (Alexander 1995). It ignores extraneous smells such as perfume. Fingerprint identification systems are now being used in retail stores and for access to ATMs (Anonymous 1996), whilst in California, a company 'Identix', has developed a system which has fingerprint recognition sensors on mobile telephones, computer keyboards, and plastic cards (Young 1999). The Bank of Texas has also recently introduced iris recognition systems for its ATM network.

The costs and volume of data required to be stored online to enable comparison for any potential user may, however, be prohibitive and there is always the possibility that computer security systems could be compromised by reproducing data streams which correspond with the biometric characteristics in question. An additional problem is that users must be required to provide samples of their personal characteristic and that the security of these samples could be compromised. Recognition systems are also, at present although not presumably in the future, costly and sometimes slow to use.

Digital Signature Security

The use of public key encryption systems also have their security risks. Public key systems require that cryptographic key pairs be issued to individuals who are able to establish their identity to an appropriate degree of assurance by supplying multiple and independent sources of identification such as those required when accounts are opened with a financial institution. Primary documentation (such as a passport, birth certificate etc) along with matching secondary documentation (such as a bank statement, car registration papers etc) would be required in order to satisfy the degree of documentary evidence of identity required.

This, however, may prove to be one of the system's weakest points in terms of security. Already systems which require the identification of individuals when they open accounts with financial institutions have been circumvented by offenders producing documents which have been forged or altered through the use of computerised desk-top publishing equipment. Birth certificates are particularly susceptible, as they can be fraudulently obtained in some jurisdictions with little difficulty by tendering scant details to the issuing authority. Birth certificates often do not entail cross-referencing, such as address, nor are they amended when the subject is deceased. Fraudulently obtained birth certificates may then be used to obtain other false documentation, such as passports and drivers' licences. Many primary documents are now protected through the use of various security devices, such as holograms, micro-printing

and void pantographs (which reveal the word 'void' when photocopied). Digitally-produced passports are also now being made with enhanced levels of security. Most of these, as well as other security devices, have all been compromised, however. Unless staff who inspect such documents are fully trained in recognising false or altered documents, it is possible to open various accounts in a variety of false names and make use of all of the banking facilities available, including loan facilities, until such time as the fraud is discovered or the false identity made known.

An example of the weaknesses of this manner of establishing one's identity arose recently in Victoria, Australia where an offender opened forty-two separate bank accounts throughout the Melbourne metropolitan region making use of false identification documents. Each account made use of a different false identity created by the offender using desk-top publishing equipment. Forty-one false birth certificates were produced along with forty-one false student identification cards, some containing photographs. Eventually, a driver's licence was obtained by relying on the false documents already produced. Along with the false bank accounts, the offender was able to register a business name, and make withdrawals from cheque accounts totalling tens of thousands of dollars. Health care refunds were also obtained and various retailers defrauded (Morton 1998). There are various solutions to the problem of counterfeit identification documentation fraud.

First, and perhaps most importantly, is the need to validate identification documents with the issuing source. Staff presented with a birth certificate should, for example, check if the details correspond with those held in the central office of Births Deaths and Marriages. Checking with the electricity company concerned should validate an electricity account tendered as an identification document. This may not always solve the problem, however, as telephone answering services can be manipulated to support the creation of false employment or identity details.

Secondly, staff involved in validating documents need to be instructed as to the security features which are present on original documents, what original documents look like, and how forged documents appear.

Thirdly, modern security features should be incorporated on all primary identification documents and even secondary documents if at all possible. Among these new technologies are security printing, in which colour-coded particles are embedded into the medium, 'tracer fibre' which can be woven into textile labels, and hidden holographic images which can be read with a hand-held laser viewer or machine reader, thus permitting verification of a product's origin and authenticity. Similarly,

technologies have already been developed which make counterfeiting extremely difficult. In any system which entails the identification of individuals to whom cryptographic keys are to be issued, these primary fraud prevention procedures would need to be included.

Once questions of identification have been resolved, issues would arise in relation to the manner in which keys or hardware tokens are given to users. Standards would also need to be complied with for the storage and use of keys, perhaps by requiring keys to be used off-line or with the use of a smartcard which is able to process transactions.

The problem remains, however, that private key data or tokens themselves must be communicated to users. The financial world has already experienced considerable problems in transferring possession of plastic payment cards to users and similar problems could arise with respect to cryptographic keys which are stored on smartcards. Adequate security precautions would need to be used to ensure that tokens are passed securely to users from the issuing authority. Another area of risk concerns the generation of cryptographic keys. It may be possible for the individual who generates a public and private key pair to retain a copy of the private key for later illegal use.

Legislation may need to be enacted which will hold the key generator liable for subsequent losses which arise out of the compromise of a key issued by that generator. Cryptographic keys would be kept on the hard drive of a computer with the cryptographic service activated by a smartcard inserted into the PC. Smartcards may also be used to sign a digital signature and to authenticate the identity of a user.

In addition to the risks associated with compromising access mechanisms such as PINs, passwords, and biometric devices, the possibility exists that smartcard tokens themselves may be altered or counterfeited. Already this has taken place in relation to smartcards used for small value commercial transactions. Where keys are stored on personal computers or servers, their security may be compromised in which case appropriate risk management measures need to be taken.

Fraud Detection Software (Neural Networks)

If one is unable to prevent on-line fraud from taking place entirely, then at least it may be possible to identify the presence of fraudulent transactions quickly in order to reduce the extent of any losses which are suffered or the occurrence of repeat victimisation.

A number of organisations are now providing software for use in the prevention of electronic funds transfer fraud. Software has been devised

to analyse user spending patterns in order to alert individuals to the presence of unauthorised transactions and also merchant deposit monitoring techniques to detect claiming patterns of corrupt merchants. Nestor Inc., for example, provides software called PRISM (Proactive Fraud Risk Management) which is used to detect credit card fraud such as lost cards, stolen cards, counterfeit cards, fraudulent applications, cards never received, mail order, phone order and catalogue sales and merchant fraud. It is designed for use by credit card issuers, credit card processors, credit card acquirers, Merchant Banks, and anyone who has over 500,000 card holder accounts. The cost is between A$300,000 and A$1,500,000 depending on system requirements and configuration (see Nestor Inc. 1996). Similar types of systems could be adapted to monitor on-line transactions. The success of such an approach depends, however, upon the extent to which the software cannot be interfered with or modified.

Improved Cryptography

Finally, technology is able to secure the content of electronic communications through the use of encryption systems. Cryptography is, however, a double-edged sword in the field of electronic commerce as it is able to protect data from unlawful interference while at the same time concealing illegal activities from lawful investigation (Denning 1995).

Nonetheless, cryptography is still employed as a mainstay of electronic banking security systems and this will continue to be the case in respect of Internet commerce

3.4 Institutional Self-Help

Financial institutions are able to adopt a wide variety of self-help strategies which may reduce the risk of funds transfer fraud on the Internet. First, they need to adopt in-house security procedures and to ensure that staff are checked for security risks and breaches, as Internet fraud will often involve confederates with inside knowledge of an institution’s security and computer procedures. If plastic card payment systems are used in conjunction with the Internet, various procedures will need to be adopted to ensure that plastic cards are not stolen and that PINs are not compromised. Financial institutions may also be able to assist Internet merchants by notifying them of stolen cards and PINs. Again in the United Kingdom, a National Hot Card File was created by which details of lost and stolen cards were quickly transmitted to retail outlets. Systems may also need to be created by which Internet merchants are able to obtain immediate authorisation from financial institutions before transactions may be accepted. It may be necessary for

all Internet transactions to be authorised before they are accepted. Card issuers are able to protect themselves against Internet fraud involving plastic cards in a variety of ways. Banks, for example, now have a centralised fraud reporting and investigation agency for plastic cards, Cardlink Services Limited, which has close liaison with the police. Cardlink Services investigates cases of fraudulent use of cards in each State and Territory of Australia and gathers evidence which is then forwarded to police (Van Rhoda 1991).

Frauds in which merchants are involved constitute a large problem for financial institutions as merchants or their employees are ideally placed to permit access to computer networks and to alter transaction details. Bonney (1992) discusses various ways in which merchants can help to prevent credit card fraud including the conduct of random authorisation checks by merchants of bank account details, merchant advertising of the fact that steps are being taken to prevent credit card fraud and closer examination of cards by sales staff when they are being used. Newton (1995) also discusses various strategies to prevent merchant abuse in relation to plastic card fraud. Similar strategies may need to be adopted to prevent Internet funds transfer fraud.

Where such strategies have been consistently implemented, substantial reductions in fraud can occur. In Britain, for example, the use of a variety of strategies designed to prevent plastic card fraud resulted in a forty-one per cent reduction in such fraud overall between 1991 and 1994, while losses occurring at retail points of sale were reduced by forty-nine per cent during the same period. Losses from cards lost or stolen in the post were reduced by sixty-two per cent between 1991 and 1994 (Webb 1996).

Codes of Practice As an alternative to the use of legislative regulatory controls, the banking and credit industries have relied on the use of codes of conduct to prevent fraud and to resolve disputes between institutions and customers. Codes have the dual function of acting as a form of education and publicity for both institutions and customers, as well as providing a statement of recommended practice which may be relied upon to resolve individual disputes.

Codes of practice are, however, only going to be an appropriate regulatory mechanism where financial institutions or system operators are involved. If electronic money or stored value cards are used on the Internet, then only the consumer and the merchant may be involved. This may require existing codes of conduct to be re-written.

In Australia, all Australian suppliers of electronic funds transfers (EFT) have agreed to comply with the Electronic Funds Transfer Code of

Conduct which was introduced in December 1989. The existing code is, however, limited to transactions involving an EFT plastic card and a PIN only, thus excluding home banking and Internet electronic money transactions which do not involve cards. The code exempts the card holder from liability in respect of fraudulent or negligent conduct on the part of card issuers' employees or agents; forged, fault, expired or cancelled cards; losses occurring prior to receipt of the card or PIN; unauthorised transactions occurring after notification; and losses resulting from unauthorised transactions where it is clear that the card holder has not contributed to the losses. The card holder's liability is limited to A$50 or the balance of the account or the loss at the time of notification of loss or theft of the card. The code, therefore, provides a wide range of rules which both institutions and users are required to follow in order to ensure that fraud is minimised and that disputes are fairly resolved.

Another code which has relevance to electronic transactions involving banks is the Code of Banking Practice introduced in November 1993. This code sets out the privacy requirements which banks are obliged to adhere to in dealing with customers, and also specifies the various rights and duties of banks and customers. This code will also be relevant to the conduct of transactions on the Internet.

Most recently, Australian building societies and credit unions have established codes of practice which regulate, amongst other things, transactions conducted electronically.

Adoption of these Codes of Practice will take place in conjunction with the introduction of Australia's Uniform Consumer Credit Legislation (Australian Payments System Council 1996). A draft Code of Conduct for the Smart Card Industry has also been prepared which deals with issues of privacy, confidentiality, disclosure and dispute resolution (reproduced in Privacy Law and Policy Reporter, 1996).

3.5 Legislative Measures and Codes of Practice

Conducting a secure system for electronic commerce may require various laws to be examined in order to ensure that fraud may effectively be investigated and prosecuted. A range of different approaches have been taken to law reform internationally in order to accommodate on-line commerce with some parliaments enacting highly specific reforms to define 'documents', 'writing', and 'signatures' as well as to specify the rules which govern the attribution of communications.

In Australia, a more generalised approach is being adopted with the enactment of broad,technology-neutral provisions which would constitute a basis for more specific legal changes which could be introduced subsequently (Australia, Attorney-General's Department 1998). As an alternative to the use of legislative regulatory controls, the banking and credit industries have relied on the use of codes of conduct to prevent fraud and to resolve disputes between institutions and customers. Codes have the dual function of acting as a form of education and publicity for both institutions and customers, as well as providing a statement of recommended practice which may be relied upon to resolve individual disputes.

Codes of conduct are, however, only going to be an appropriate regulatory mechanism where financial institutions or system operators are involved. If electronic money or stored value cards are used on the Internet, then only the consumer and the merchant may be involved. This may require existing codes of conduct to be re-written.

It may now be appropriate for an Internet Commercial Code of Conduct to be established in order to deal with the allocation of risk and determination of liability involved in Internet-based transactions. Issues which could be dealt with in the Code could include: guidelines on users' obligations in maintaining computer hardware in a secure environment; principles to be observed for obtaining, storing, and using encryption keys securely; principles to be observed for storing keys securely and for preventing unauthorised access to tokens; obligations to be complied with regarding security and privacy of data; and principles to be observed in determining liability and the allocation of loss arising out of the use of the use of the Internet.

4.0 CONCLUSION

This unit has described a wide variety of ways in which funds may be stolen through exploiting security flaws in electronic funds transfer systems used in conjunction with on-line commerce. The range of electronic systems used to conduct commercial transactions is increasing rapidly and considerable effort is being directed at ensuring the security of digital transmissions which represent monetary value. The opportunities for fraud are, however, substantial. The solution to electronic funds transfer crime on the Internet will ultimately involve the adoption of a range of strategies both technological and strategic in which close cooperation will exist between all those involved in providing and using systems. This includes telecommunications carriers and service providers, financial institutions, retail merchants, and individual users.

One area of particular importance relates to the need for strategies to be used which will enable the emergence of weaknesses in systems to be quickly identified. Once recognized, there should be a prompt response to the problem. In ensuring that particular weak points in security systems are identified and weaknesses solved, it is likely that technology will provide the most effective response.

Probably the greatest source of risk in conducting on-line business lies in the area of user authentication. False identity fraud has been a continuing problem in commerce for decades now and it is likely that it will continue in adapted forms on the Internet. If passwords continue to be used to restrict access to computers then they should be protected by the various security devices I have mentioned. Biometric identifiers will, presumably, become much more widely accepted as Internet commerce develops.

In planning for the future, it will be necessary to ensure that the weak points in security protocols are not overlooked. As in other areas of fraud control, the weak points in on-line commerce will invariably arise out of human factors rather than technological considerations.

5.0 SUMMARY

- The extent to which individuals are using the Internet for business transactions is increasing enormously. The market is potentially great with approximately 171.25 million people estimated to be using the Internet worldwide in May 1999 (NUA Internet Surveys 1999a).
- One of the most effective strategies used to control crime is education of the public as to the nature of the security risks which they face, and how they may best protect themselves.
- Institutions involved in maintaining the infrastructure of the Internet as well as financial institutions are able to adopt a wide variety of self-help strategies which may reduce the risk of funds transfer fraud on the Internet.
- A wide range of technological solutions have been devised in order to reduce the security risks associated with conducting on-line business.
- Financial institutions are able to adopt a wide variety of self-help strategies which may reduce the risk of funds transfer fraud on the Internet.
- Conducting a secure system for electronic commerce may require various laws to be examined in order to ensure that fraud may effectively be investigated and prosecuted.

6.0 TUTOR-MARKED ASSIGNMENT

Briefly discuss Cardholder Verification as a technical measure in electronic fraud prevention.

7.0 REFERENCES/FURTHER READINGS

Alexander, M. (1995). The Underground Guide to Computer Security. New York: Addison-Wesley Longman Inc.

Anonymous (1996). "Fingerscan's $2.5m Deal", Security Australia, vol. 16, no. 10, p. 2.

Australia, Attorney-General's Department (1998). Electronic *Commerce: Building the Legal Framework.*

Report of the Electronic Commerce Expert Group to the Attorney-General, Australian Government Publishing Service, Canberra.

Bellcore (1996). "New Crypto-Attack Weakens Seeming Strength in Smart Cards, Secure ID Cards, or Vale Cards", Internet http://www.infowar.com/sample/infosec4.html-ssi

Bowes, C, (1996). "Digital Dollars", Bulletin, 20 August: 50.

Carter, S, (1996). "Online 'Bank' Cashes in on Cyber Commerce", The *Australian, 30 July, Computers, p. 49.*

Da Silva, W. (1996). "'Hackers' May Evade Charges", The Age *(Melbourne), 11 June: C1.*

Denning, D. (1998). 'Cyberspace Attacks and Countermeasures', in Denning, D. E. and

Denning, P. J. Internet Besieged: Countering Cyberspace Scofflaws, ACM Press, New York, pp. 29-55.

Denning, D. E. (1999). Information Warfare and Security, ACM Press, Reading, Massachusetts.

Hansell, S. (1996). "AT&T and Wells Fargo Investing in an Electronic Cash Card", New York Times, 19 July, p. C2.

Holland, K. (1995). "Bank Fraud, The Old-Fashioned Way", Business *Week, 4 September, p. 88.*

Johnson, E. (1996). "Body of Evidence: How Biometric Technology Could Help in the Fight Against Crime", Crime Prevention News, December, pp. 17-19.

Jupiter Communications (1999). 'Travel Suppliers Missing Online Market Potential',

Levi, M. and Handley, J. (1998). The Prevention of Plastic and Cheque *Fraud Revisited, Home Office Research Study No. 182, Home* Office, London.

Levy, S. (1994). "E-Money (That's What I Want)", Wired, December, pp. 174-9, 213.

Mackenzie, R. (1998). "Virtual Money, Vanishing Law: Dematerialization in Electronic Funds Transfer, Financial Wrongs and Doctrinal Makeshifts in English Legal Structures", *Journal of Money Laundering Control, vol. 2, no. 1, pp. 22-32.*

Mackrell, N. (1996). "Economic Consequences of Money Laundering", in Graycar, A. and

Grabosky, P. (eds.), Money Laundering in the 21st Century: Risks and *Countermeasures, pp. 29-35, Australian Institute of Criminology,* Canberra.

Meijboom, A. P. (1988). "Problems Related to the Use of EFT and Teleshopping Systems by the Consumer", in Poullet, Y. and Vandenberghe, G. P. V. Telebanking, Teleshopping and the Law, Kluwer Law and Taxation Publishers, Deventer, pp. 23-32.

Morton, G. (1998). Personal communication, Detective Sergeant Gavin Morton, Victoria Police, Major Fraud Group, 22 January 1998.

Nestor Inc. (1996). "Proactive Fraud Risk Management: Neural Network Based Credit Card Fraud Detection from Nestor Inc." Internet

NetCrusader (1998). 'NetCrusader Product Family: Security Solutions for the Enterprise',

Newton, J. (1995). Organised Plastic Counterfeiting, HMSO, London.

Nicholson, E., 'Hacking away at liberty', Times (London), 18 April 1989.

NUA Internet Surveys (1999a). 'Internet Usage',

NUA Internet Surveys (1999b). 'National Consumers League: Seven Percent of US Users Hit by Credit Card Fraud',

Rawitch, R. (1979). "Expected Bank Plot to Fail", Los Angeles Times, 23 February, pp. 1, 27.

Smith, R.G.. "The Prevention of On-Line Financial Fraud", Research Analyst, Australian Institute of Criminology

Smith, R. G. (1997). "Plastic Card Fraud", in Trends and Issues in *Crime and Criminal Justice, No. 71, Australian Institute of* Criminology, Canberra.

Spinks, P. (1996). "Tests Show Up Smart Card Flaws", The Age *(Melbourne), 6 December.*

Sullivan, C. (1987). Unauthorised Automatic Teller Machine Transactions: Consequences for Customers of Financial Institutions", Australian Business Law Review, vol. 15, no. 3, pp. 187-214.

Visa International (1997). "SET Draft Reference Implementation", International Society for the Reform of Criminal Law

Webb, B. (1996). "Preventing Plastic Card Fraud in the UK", Security *Journal, vol. 7, pp. 23-5.*

Young, S. (1999). 'Thumbs Up for Fingerprint-Based Ids', The Age *(Melbourne), IT p. 4.*

UNIT 4 SANCTIONS AGAINST PLASTIC CARD FRAUD: THE CASE OF AUSTRALIA

CONTENTS

1.0 INTRODUCTION

Plastic card fraud in the twenty-first century is no longer a domestic problem for financial institutions, merchants, consumers and law enforcement agencies but is now of global concern.

Developments in payment systems enable transactions to be carried out with ease across borders and plastic cards are one of the primary means of gaining access to accounts regardless of where the customer is located. In Europe, for example, the relaxation of borders and the introduction of the Euro have led to a substantial increase in cross-border transactions, and associated fraud. Similarly, developments in electronic commerce have meant that a much higher proportion of transactions are carried out without cards actually being present.

By 2005, MasterCard international estimates that 52 per cent of its transactions will be carried out through remote services, rather than at point of sale (Lisker 2001). In Europe, the proportion of fraud on United Kingdom- issued plastic cards committed outside the United Kingdom, doubled during the 1990s and will soon amount to one third of all losses (Levi 2000). Fraudsters are clearly making use of the many new opportunities that have been created through the increase in international card based transactions and the corresponding weaknesses in fraud control measures that they entail. KPMGs latest Fraud Survey (2002), for example, noted an increase in the number of international criminals coming into Australia and New Zealand, committing major fraud and then leaving with the proceeds of their crimes. The report also found an increase in the involvement of criminal gangs in external fraudulent attacks on financial institutions through the use of stolen cheques and falsified identification documents. Of the 148 respondents with international operations, 23 per cent had experienced fraud in the preceding two years involving A$30 million.

The most prevalent types of fraud within respondents' off-shore operations were credit card fraud and theft of inventory.

In the European Union, plastic card counterfeiting is estimated to cost EUR600 million or 0.07 per cent of industry turnover (Lakeman 2001), while much of the increase in plastic card fraud has related to card not present transactions conducted by telephone and the Internet.

Unlike crimes involving personal violence in which the offender and the victim have to be present together in one place at one time, economic criminals and their victims can be located anywhere in the world—and sometimes never meet in person. Occasionally, the offender and the victim may be located in one jurisdiction, but the mechanics of the commission of the offence may entail an international component, or involve confederates in a third country. This gives rise to various legal problems in determining exactly where the offence occurred, which country has jurisdiction to deal with it, as well as problems for financial crime investigators and police of locating offenders, obtaining evidence (often in foreign languages), and in seeking extradition of offenders.

It is also important to recall that there may be many different types of offenders involved in plastic card fraud. These include street robbers at the lowest level of organizational networks who steal cards; skilled offenders who have the technological expertise to skim, counterfeit, and alter cards; offenders who work within organisations who are able to steal account information for use by others; dishonest merchants who misuse customer account information; offenders who use counterfeit cards to withdraw cash or to obtain goods and services illegally; and those who control and organize the activities of other actors in the illegal pipeline. Each of these categories of offenders operates in slightly different ways, is subject to different motivations and will respond differently to legal sanctions that might be imposed. To speak of a singular category of plastic card fraudster is, therefore, inappropriate.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- identify the facts about plastic card fraud as well as the benefits
- understand what motivates plastic card fraudsters
- explain to what extent jurisdictions can go in deterring plastic card frauds
- understand the International Convention and Articles against plastic card frauds
- use Australia as a case to understand legislative responses in dealing with e-frauds.

3.0 MAIN CONTENT

3.1 Rational Choice Offenders

The difficulties associated with prosecuting cross-border plastic card fraud are such that hard decisions have to be made about whether or not to embark on a prosecution in the first place. If relatively small sums are involved it may well not be cost-effective to expend further funds in investigating and in prosecuting the matter.

This raises the difficult problem of when it is appropriate to take criminal proceedings and whether the results obtained, in terms of conviction and punishment, warrant the costs involved.

In 1764, the Italian criminologist, Cesare Beccaria described the purposes of punishment as being 'to dissuade the criminal from doing fresh harm to his compatriots and to keep other people from doing the same' (Young 1986, p. 23). In any discussion of the effectiveness of punishment, it is important to distinguish between these two forms of deterrence, namely that which is directed at the individual offender (special deterrence) and that which applies to other members of the community (general deterrence).

The primary motivation of fraudsters now, as in Beccaria's time, remains the same, namely cupidity in obtaining money through unlawful means. The explanations offered also remain the same: greed in supporting a particular lifestyle, necessity in feeding an addiction to drugs or gambling, envy, or curiosity in seeing how far security systems can be compromised.

There is clear evidence that the courts and the community rely upon the imposition of criminal sanctions as being a deterrent to crime, both individually and generally (see Flanagan and Longmire 1996). Indeed, deterrence is legislatively specified as one of the purposes for which sentences may be imposed by courts in various jurisdictions (e.g. *Sentencing Act 1991 (Vic.) s. 5(1)(a)), and is often relied upon by courts* as a justification for imposing a sentence of imprisonment for serious offences (see Fox and Freiberg 1999).

In determining whether legal sanctions have any deterrent effect, consideration needs to be given to a number of factors. First, whether offenders are influenced by matters which are not directly relevant to their offending behaviour, such as the likelihood of detection, arrest and a particular punishment being imposed, secondly, whether offenders are, in fact, aware of the probability of detection, arrest and particular

punishments being used or know of sentencing policies and practices generally, and thirdly, whether they are minded to act upon any such knowledge by modifying their behaviour or propensity to commit crime. Serious doubts have been expressed concerning each of these matters.

A large body of research has sought to test the so-called rational choice perspective of offending, namely, whether offenders weigh up in advance of committing crimes the positive and negative factors relevant to a course of criminal conduct, and then make a rational decision that seeks to maximise their net gains (Clarke and Felson 1993).

An example of the kind of calculations that plastic card offenders might make when deciding to commit either plastic card fraud or armed robbery was described by John Newton who undertook an extensive study of organised plastic card counterfeiting in 1994.

> The attractions of plastic fraud to the criminal are clear. It is a low-risk, highly profitable venture. . . . There is no need to obtain firearms to commit the offence. There is little chance of being caught in the act and absolutely no chance of being shot by police, which is a hazard of committing an armed robbery in any country. . . . If criminals are caught it is difficult and costly to prove the case against them in court. If there is a suitable criminal offence in the country concerned . . . which is by no means certain, and criminals are convicted, the chances of being sent to prison are about even. If they are sent to prison, their sentence will be substantially less than one they would receive for armed robbery (Newton 1995, p. 101).

The decision- making employed by offenders relates to all aspects of criminal justice processing—namely, whether their crime will be detected by investigators, whether they will be arrested, extradited, granted bail, required to undergo a jury trial, convicted, what sentence will be imposed, and how that sentence will be carried out, for example, in a maximum security city prison or in a low security institution, and at what point parole will be obtained.

Offenders may also weigh up rationally the likelihood of other consequences of their wrongdoing occurring—namely, whether they will lose their job, lose professional registration, be barred from being a company director, be unable to obtain work on release from prison, be forced to sell their home, suffer marital problems, health problems, and loss of respect amongst family, friends, and peers.

Research into the rational-choice perspective of offending has, however, produced largely inconclusive findings. Surveys of offenders have sought to identify the matters taken into account at the time the crime

was committed (see Cornish and Clarke 1986) while other studies have sought to assess the impact of sanctions on rates of recidivism (Weisburd and Waring 2001, p. 150). Both have found that offenders do not generally take the risk of punishment into account when deciding whether or not to commit crimes.

In Western Australia, for example, Harding (1990) surveyed 469 prisoners who were serving sentences for violent crimes and found that some knew of the penalties associated with using a firearm in a robbery and made a rational choice to carry out the offence nonetheless.

Unfortunately, ninety-one per cent of offenders who carried a firearm in robberies said that they would carry a gun the next time they committed a robbery-thus providing little deterrent effect of imprisonment.

A study in the United States that examined the motivations of offenders who carried out serious property crimes, involved a survey of sixty offenders who were serving at least their second term of incarceration for offences such as burglary and armed robbery (Tunnell 1996). All sixty respondents in the study reported that they and nearly every thief they had ever known simply did not think about possible legal consequences of their actions. Although the offenders knew that they were doing wrong and tried to avoid arrest, thirty-two did not know the penalty attaching to their act until after their arrest. Thirty-six of the respondents said that the possibility of incarceration was no threat to them and the remainder did not perceive it as being a great threat. Thus, incarceration could not be said to have acted as a deterrent to the majority of these serious repeat property offenders.

In a study of official re-offending rates amongst a sample of white collar offenders sentenced by federal courts in the United States, a similar finding was obtained, namely that imprisonment did not influence the likelihood of re-offending for those convicted of white collar crimes. Prison sentences did not have a specific deterrent effect on re-arrest whether in terms of the likelihood, timing, frequency or type of recidivism (Weisburd and Waring 2001, p. 113). Some groups of offenders may be particularly unlikely to be influenced in their offending behaviour by the possibility either of detection or the threat of subsequent incarceration. For example, where crime is committed out of extreme need, under duress, through the influence of alcohol or drugs, or because of an addiction to gambling, it is unlikely that such offenders will have regard to the possibility of arrest and punishment when determining whether or not to carry out the crime in question (Fox and Freiberg 1999). For them, rational decision-making regarding the possibility and nature of punishment is unlikely to be present at the time of offending. It is, therefore, important to consider the individual

circumstances of the offence and the offender when assessing any potential deterrent effects.

The extensive criminological research on the effects of deterrence is inconclusive regarding the influence which different types of sanctions have with respect to the prevention of crime. Even incapacitation (that is, removing offenders from the community by keeping them in detention) may be of little importance in reducing crime as others in the criminal community may simply take the place of those incarcerated (see Chan 1995, p. 10). The courts, legislators and the public, however, generally believe that the possibility of incarceration being imposed does act as a deterrent to crime. This is particularly the case with respect to general deterrent effects, although less so with respect to marginal deterrent effects, that is, the effect which increasing a penalty has on the likelihood that offenders will be deterred more effectively from committing that offence (Zimring and Hawkins 1973).

The Victorian Sentencing Committee, which conducted a thorough review of the research available at the time, arrived at the following conclusions with regard to the deterrent effect of increasing penalties (Victoria, Attorney-General's Department 1988, p. 77):

> It may be accepted on the basis of the available evidence and common sense that the existence of a penalty system in force through the criminal justice system will result in a general deterrent effect in the commission of crime. . . There is no evidence to support the proposition that there is any marginal deterrent effect in either increasing a specific penalty imposed on a given offender; [or] increasing by legislative means the general level of penalties applying for a given offence. . . There is no means of accurately assessing any marginal deterrent effect that may exist in given situations.

In the case of organized criminal activity, such as that involving cross-border plastic card fraud, rational choice may play a greater role than in other types of violent or property crime. Large-scale counterfeiting operations require planning and consideration of the costs and benefits including the possibility of punishment being imposed. Organised criminals who embark upon cross-border crime, however, can be fairly certain that prosecution will be difficult and costly for the authorities and thus, that they may well escape punishment. Some organized groups plan their activities carefully so that the risk of detection is low, that those responsible for organizing operations are less likely to be arrested, and that prosecution will be difficult even if arrest does occur. They do this by moving operations regularly, changing modus operandi, making use of false identities, and perpetrating fraud on many occasions, each involving relatively small sums.

This somewhat pessimistic view about the effectiveness of criminal sanctions may lead to the conclusion that it unnecessary for victims to report crimes to the police at all. It is important to recall, however, that if crimes are not reported, not investigated and do not result in punishment, then any deterrent effects—however small—are likely to be diluted and those individuals who do rationally consider the consequences before offending, will be more likely to see the possibility of punishment as remote, and thus more likely to offend.

In order to prevent and to deter crimes of this nature, therefore, there is a need for as many cases as possible to be dealt with formally and for judicial outcomes and other consequences of wrongdoing to be widely publicised.

3.2 Jurisdiction Shopping

One consideration that rational-choice offenders may take into account in deciding whether or not to offend is the likelihood that they will be prosecuted. In the case of cross-border crimes where the offender is located in a different jurisdiction from the victim, a rational-choice offender may select a jurisdiction that will minimise the chances of prosecution. Difficulties in investigation and prosecution can occur because of policies of bank secrecy in particular countries that make it difficult to obtain evidence from financial institutions, because mutual assistance arrangements do not exist between the countries in which the offender and victim are located, because of an unwillingness or inability on the part of police to investigate these types of crime, or the absence of extradition treaties with the country in which the offender is located when arrested. Other impediments to prosecution include the cost of sending law enforcement officers abroad to assist in an investigation, and the cost of bringing witnesses from abroad to testify in proceedings, which may both be prohibitive. There may also be problems of language, geographical distance, lack of knowledge of foreign le gal systems, time differences, telecommunications and technological differences, and expense. These are all substantial impediments to embarking on a prosecution.

There may also be significant legal impediments which must be overcome. Some countries do not have laws that proscribe the possession of counterfeit cards or card embossing machines, and offenders may choose these countries in which to base their operations (Newton 1995, p. 38).

Other offenders may continually move their operations in order to make detection difficult. The laws of evidence may also make evidence

obtained in one country unable to be used in criminal proceedings in another country. These problems are not, however, new and international law has had to cope with the complexities of jurisdictional issues and conflicting substantive and procedural laws for hundreds of years in prosecutions involving sea piracy, slavery, hijacking, war crimes, and other offences that have an international component.

The result is that some offenders may be able to commit their crimes with relative impunity and be unable to be dealt with.

Those few fraudsters who think rationally about the consequences of offending could also target victims in countries that have the lowest maximum penalties for relevant offences or those in which sentencing practices result in comparatively low terms of imprisonment being imposed for the types of offences being contemplated.

In the field of money laundering, these notions have been clearly demonstrated with some nations being seen as safe jurisdictions in which to base criminal activities. In the case of plastic card fraud, this is also likely to be the case, at least with respect to large scale, organised activities. Offenders in China, for example, where the death penalty exists for serious fraud offences, would clearly be well-advised to target victims in Australia where, in some states, they would receive a few years' imprisonment, or less, for offending, if they were able to be prosecuted at all.

The main responses to jurisdiction shopping by offenders are, first, sharing of information between law enforcement and regulatory agencies and, secondly, harmonisation of laws internationally.

In terms of cooperation between agencies, in July 2000, an important initiative began when the United States Federal Trade Commission (FTC) entered into an agreement with the Australian Competition and Consumer Commission to provide access to the FTCs Consumer *Sentinel database of consumer complaints. This now permits regulators* in the United States, Canada, and Australia to share information about consumer complaints and to assist each other in crossborder prosecutions—such as those involving Internet sales and on- line auctions. Information networks within and between law enforcement agencies also need to be used, so that when an investigation begins, contact can be made immediately with the appropriate person in another country's corresponding department. Secure Intranets, such as that used by the Australian Bureau of Criminal Intelligence, are an excellent way in which this can be achieved.

They can also be used to share 'Fraud Alert' information and to exchange intelligence needed in investigations. Twenty- four hour response centres are now being established in many countries. These centres, which are to be used for genuine emergencies only, enable requests for real-time computer investigations to be handled at any time of the day or night in the participating country. In Australia, the Australian Federal Police handles such requests and refers queries to relevant state and territory police services of other Australian Federal Police regional offices (Geurts 2000). In another initiative in the United States, the Federal Bureau of Investigation and the National White Collar Crime Centre have co-sponsored the establishment of a central repository for complaints relating to Internet fraud. The Internet Fraud Complaint Centre (IFCC) hopes to ensure that Internet fraud is able to be addressed at all levels of law enforcement (local, state, and federal)

The IFCC was created to identify, to track, and to investigate new fraudulent schemes on the Internet on a national and international level. IFCC personnel collect, analyse, evaluate, and disseminate Internet fraud complaints to the appropriate law enforcement agency. The IFCC provides a mechanism by which Internet fraud schemes are identified and addressed through a criminal investigative effort. The IFCC also provides analytical support, and aid in the development of training modules to address Internet fraud. The information obtained from the data collected provides the foundation for the development of a national strategic plan to address internet fraud.

In the European Union, Europol, which was created in 1998 and based in the Hague, is an information clearing house and analysis centre with law enforcement liaison officers in various member states. It aims to increase cooperation and communication between and among law enforcement agencies in member states rather than acting as a European police service (Sussmann 1999, p. 480).

In 1996, the G-8 countries established a group of experts ('The Lyon Group') to examine better ways in which to fight international crime. The Group produced forty recommendations that were endorsed by the G-8 heads of state at the Lyon Summit in June 1996. This group has met regularly and has discussed ways of enhancing the ability of law enforcement agencies to investigate and prosecute international crime. In January 1997 it created a sub-group to look specifically at high-technology crime and this sub- group has examined law reform, investigatory, and procedural issues to do with prosecuting cross-border computer crime (Sussmann 1999). The G-8s High-Tech Crime Group, as it is known, has also recommended the establishment of cooperative arrangements between public sector police and regulatory agencies and the private sector. For example, there is a need for telecommunications

carriers and ISPs to make certain information available to investigators on production of an appropriate search warrant. Ideally, such arrangements need to be uniform across jurisdictions. One example of a cooperative venture involving public and private sector bodies is the Cybercrime Unit created by the International Chamber of Commerce's Commercial Crime Bureau in London in 1999. This brings together law enforcement bodies such as Interpol, Scotland Yard, and the FBI, as well organisations within the private sector including major financial institutions and businesses. The Unit acts as a clearinghouse for information on electronic crime and passes details of frauds and solutions between companies and the police. Cooperative cross-border ventures to deal with money laundering have also been established. The International Money Laundering Information Network is an Internet-based network assisting governments, organisations and individuals in the fight against money laundering. IMoLIN has been developed with the cooperation of the world's leading anti- money laundering organisations that include the Commonwealth Secretariat, the Council of Europe, the Financial Action Task Force, Interpol, the United Nations Office for Drug Control and Crime Prevention's Global Program against Money Laundering, the European Commission, and others. The Egmont Group of the Financial Action Task Force also coordinates the activities of various Financial Intelligence Units globally.

3.3 International Conventions

Turning to the question of harmonisation of laws, we have also seen some important developments take place in recent years that are relevant to the prosecution of financial crimes. International conventions need to deal not only with substantive laws relating to crimes of dishonesty, but also jurisdictional and procedural laws concerning mutual legal assistance. In particular laws concerning search and seizure need to be consistent and complementary internationally so that police can obtain evidence from other jurisdictions. Law reform is, however, essentially a matter for each individual nation. As the Lord Chancellor, Lord Halsbury observed in 1891 in the case of Macleod v Attorney-General *of New South Wales ([1891] AC 455, 458), 'all crime is local'. This does* not mean, however, that parliaments should reform laws in total disregard of reforms introduced elsewhere. In the case of cross-border financial crime, all aspects of the judicial process would be facilitated if as much uniformity as possible were introduced in relevant laws. This would prevent jurisdiction shopping and would also enhance uniformity of sanctioning and reduce some of evidentiary difficulties that arise in proceedings.

Achieving uniformity of legislation is, however, neither simple nor quick. In a survey carried out by McConnell International (2000), the

laws in 52 countries were examined. Of the countries surveyed, only thirteen (25%) had up-dated their laws relating to computer-related fraud (including Australia).

The creation of multilateral treaties is also not without problems. The Council of Europe's Convention on Cybercrime (2002) took almost five years to appear. It was adopted by the Committee of Ministers of the Council of Europe on 8 November 2001 and opened for signature on 23 November 2001 in Budapest. The Convention is the first international treaty to address criminal law and procedural aspects of various types of criminal behaviour directed against computer systems, networks, or data and other types of similar misuse. As such it provides a framework for international reform in this area.

Some of the articles in the Convention that are relevant to cross-border plastic card fraud include:

3.4 Articles

Article 6—Misuse of devices

1. Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right:

(a) The production, sale, procurement for use, import, distribution or otherwise making available of:

(i) a device, including a computer program, designed or adapted primarily for the purpose of committing any of the offences established in accordance with the above Articles 2 through 5;
(ii) a computer password, access code, or similar data by which the whole or any part of a computer system is capable of being accessed, with intent that it be used for the purpose of committing any of the offences established in Articles 2 through 5; and

(b) The possession of an item referred to in paragraphs (a)(i) or (ii) above, with intent that it be used for the purpose of committing any of the offences established in Articles 2 through 5. A Party may require by law that a number of such items be possessed before criminal liability attaches.

Article 7—Computer-related forgery

Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, the input, alteration, deletion, or suppression of computer data, resulting in inauthentic data with the intent that it be considered or acted upon for legal purposes as if it were authentic, regardless whether or not the data is directly readable and intelligible.

A Party may require an intent to defraud, or similar dishonest intent, before criminal liability attaches.

Article 8—Computer-related fraud

Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, the causing of a loss of property to another person by:

(a) Any input, alteration, deletion or suppression of computer data;

(b) Any interference with the functioning of a computer system, with fraudulent or dishonest intent of procuring, without right, an economic benefit for oneself or for another person.

On 15 November 2000, another milestone was achieved with the adoption by the United Nations of the Convention Against Transnational Organised Crime. The Convention is intended to provide a legal framework for concerted action against organised crime, and the basis for the harmonisation of national legislation. It contains provisions requiring the criminalisation of certain conduct (including participation in an organised criminal group, money laundering and corruption), as well as provisions on corporate liability, special investigative techniques, witness and victim protection, cooperation between law enforcement authorities, exchange of information on organised crime, training and technical assistance, and prevention at the national and international levels. To date 141 countries have signed to Convention, including Australia which signed on 13 December 2000 at Palermo. The Convention offers great potential for enhanced cooperation among countries with respect to implementation of anti- money laundering measures, confiscation of criminal assets, promotion of extradition and mutual legal assistance mechanisms, and the application of modern technology in the fight against crime.

It is important for as many countries as possible to ratify these conventions in order for safe havens for criminals to be removed and for

prosecution and punishment of crime that takes place across jurisdictional borders to be enhanced.

Allied to the harmonisation of laws, is the need to harmonise other aspects of business practices in order to provide a global environment in which economic crime is difficult to perpetrate and yet simple to detect. Bodies such as the International Accounting Standards Committee (IASC), for example, help to promote uniform accounting practices and procedures within the business community that seek to reduce the risk of improper conduct being engaged in. Similarly, international professional bodies have roles to play in creating uniform ethical practices globally that militate against fraud (Braithwaite and Drahos 2000, p. 121).

3.5 Experience of Australia's Legislative Response

In Australia, a package of measures was adopted in the later 1980s to facilitate the prosecution of organised crime and serious fraud. These included the Mutual Assistance in Criminal Matters Act 1987 (Cth) which established mechanisms to facilitate international cooperation between investigators with respect to obtaining evidence, the location of witnesses and suspects, the execution of search and seizure warrants, the service of documents, the forfeiture of property and recovery of fines and various other matters; the Proceeds of Crime Act 1987 (Cth) which enables investigators to follow the trail of the illegal proceeds of crime internationally and to confiscate 10 assets (with a new Proceeds of *Crime Bill currently before Federal Parliament including civil forfeiture* provisions); the Financial Transaction Reports Act 1988 (Cth), establishes a government agency to monitor the movement of large-scale cash transactions and regulates the manner in which accounts with financial institutions are created; the Extradition Act 1988 (Cth) which extended Australia's ability to enter into extradition arrangements internationally; and the Telecommunications (Interception) Amendment *Act 1987 (Cth) which extended the ability of agencies to undertake* electronic surveillance for law enforcement purposes; and the National *Crime Authority Act 1984 (Cth) which created a law enforcement* agency to deal with serious and organised crime.

At present in Australia each jurisdiction has its own laws and rules governing dishonesty. Examples include theft, obtaining a financial advantage by deception, making a false instrument, fraudulent misappropriation, obtaining money by false or misleading statements, obtaining credit by fraud, false pretences, fraudulent personation, forging and uttering, using a false instrument and many others. Many of these laws are complex, unclear and contradictory and impede the successful investigation and prosecution of fraud, particularly that which takes place across jurisdictional borders (see Lanham 1997). The

Commonwealth criminal law relating to economic crime has recently been amended by the Criminal Code Amendment (Theft, Fraud, Bribery and Related Offences) Act 2000 (Cth). Various offences are specified including the offence of obtaining property or a financial advantage by deception (Division 134) and certain other offences involving fraudulent conduct (Division 135). This legislation also amends the law governing geographical jurisdiction to facilitate the prosecution of cross-border criminal activity.

Legislation was enacted on 25 November 1999 to accommodate electronic transactions with the Commonwealth. The Electronic Transactions Act 1999 (Cth) provides broad, technology- neutral provisions which constitute a basis for more specific legal changes which will be introduced subsequently. Although Australia is not a party to the European Convention on Cybercrime, it has enacted many of its articles in the form of the Cybercrime Act 2001 (Cth.) which was assented to on 1 October 2001 and commenced on 21 December 2001. This Act inserts a new Part relating to Computer Offences into the Commonwealth Criminal Code Act 1995 and provides model laws concerning a range of computer-related offences and jurisdiction that should address many of the problems that arise in prosecuting crimes of dishonesty committed electronically. The substantive offences in the Cybercrime Act 2001, however, only apply to situations in which a 'telecommunication system' or 'Commonwealth computer' is involved. Although limited in its Commonwealth focus, the Act significantly improves the scope for prosecuting cyber criminals by introducing substantive offences and procedural provisions consistent with those of the Convention. Whilst a significant improvement, there still remains some uncertainty in relation to the scope of a warrant; the ability to intercept e- mails prior to delivery; obtaining data not on premises; extra-territorial searches; and mutual assistance orders (see Ghosh 2002).

The advent of computer-related crime has meant that a number of countries are now enacting laws with extra-territorial effect—so as to permit prosecutions in respect of crimes and offenders located in some other jurisdiction as long as there is some connection with the country that enacted the law.

In terms of procedural reform, a number of improvements could be made. These include: taking early steps to ensure that evidence and facts are agreed and admitted wherever possible; streamlining interviewing procedures and using teleconferencing technologies for interviewing; using documentary evidence in preference to oral testimony wherever possible; overcoming the barriers to the use of computer-generated evidence; ensuring that evidence is not altered or destroyed before it is

able to be obtained from another country; ensuring that police have access to the plain text version of encrypted files, either by requiring the suspect to disclose the encryption key, or by employing trusted third parties to hold copies of private encryption keys which can then be used by law enforcement on production of a warrant. An additional problem is that the current multiplicity of rules makes difficult the task of effective communication of the content of those rules to business professionals and members of the public alike. As has been argued in England, reform of the substantive law regarding fraud 'could radically reduce the time and money spent on trials; increase successful prosecutions, thereby deterring many would-be fraudsters; and be a coherent foundation for preventive regulation' (Page 1997, p. 30).

3.6 Sanctions

In Australia, the following judicial punishments are currently available in most jurisdictions: fines; restitution and compensation orders; forfeiture and disqualification (confiscation); unsupervised release (suspended, deferred, conditional sentences); supervised release (probation, community service, intensive corrections) and custodial orders (either full time or periodic). In other countries, more extreme sanctions exist such as mutilation and capital punishment. In Jiddah, Saudi Arabia, for example, in May 2000, the penal authorities beheaded seven Nigerians and cut off the right hands and left feet of three others who committed an armed bank robbery (Associated Press 2000). In China, people continue to be sentenced to death for a variety of non-violent economic crimes ranging from tax evasion and value added tax fraud, counterfeiting, embezzlement to credit card theft. For example, in March 1997, Wang Hua was given a death sentence with a two year reprieve for alleged credit card theft of US$62,650. In Yunnan province, on 24 December 1997, Yang Weixiang was executed for allegedly embezzling US$72,289 from the bank where he worked (see Amnesty International 1998).

I am not recommending that Australia should introduce the death penalty for credit card fraud, but it could explore alternative sanctions. In addition to judicial punishments, there are other consequences of wrongdoing which may be invoked: adverse publicity; professional disciplinary sanctions; civil recovery action; injunctive orders and, most recently, various forms of reconciliation or community conferencing which are being evaluated at present. Judicial punishments have been described as operating within an enforcement pyramid in which the most severe penalties, which are seldom used, sit at the top of the pyramid, whilst the least severe penalties, which are frequently used, fall near the base of the pyramid. Other non-judicial regulatory responses such as warnings form the base of the pyramid in that they are used most often

(see Ayres & Braithwaite 1992, p. 35). The perceived severity of individual sanctions depends, however, upon the individual circumstances of the offender. Disqualification as a company director, may, for example, be a far more effective sanction to impose on a dishonesty than a severe fine. Similarly, adverse publicity can have profound effects in terms of shaming an offender in the community, perhaps more so than undertaking anonymous community service.

It has been argued that compliance with laws is best able to be achieved where the most severe forms of punishment, such as incarceration, are available but seldom used. In the words of Ayres and Braithwaite, 'the more sanctions can be kept in the background, the more regulation can be transacted through moral suasion, the more effective regulation will be' (1992, p. 19). The maximum penalties which attach to serious financial crimes already reflect the seriousness of such conduct with lengthy terms of imprisonment and substantial fines being available. Little research has been carried out in Australia on the manner in which white collar offenders are dealt with following a criminal trial. In a study undertaken of a sample of fifty completed cases handled by the Major Fraud Group of the Victoria Police between January 1990 and October 1994, it was found that 68 per cent of offenders were sentenced to terms of imprisonment, usually less than five years, 14 per cent received good behaviour bonds, 11 per cent received suspended terms of imprisonment, 4 per cent were fined and 3 per cent received community-based orders (Krambia-Kapardis 2001, p. 100). These cases included, however, some of the most serious fraud offences prosecuted in Victoria. It has been argued that white collar offenders tend to receive non-custodial sentences more often than custodial sentences owing to the fact that they are often first-time offenders, have cooperated with the police, have made financial restitution for their offences, may have suffered other consequences of their wrong-doing such as professional disqualification, and are invariably proficiently represented by senior legal practitioners who are able to describe their clients' circumstances in the most favourable light to the judge. Some, such as mitigating Alan Bond, were previously persons of high standing in the community.

Research supports the view that it is not the type of sentence which determines an offender's future criminal career, but rather various social and personal factors including access to employment and family and community support. Recidivism rates for offenders who have received community-based penalties, for example, do not significantly differ from recidivism rates for offenders who have experienced incarceration. Recidivism rates tend, however, to be higher for offenders who have been sentenced for more serious offences regardless of the type of sanction received, while offenders who have undergone community-based penalties suffer fewer adverse effects from the experience, which so

may cause them to re-offend, than offenders who have undergone incarceration.

It is widely known that increasing the certainty of detection through more effective and efficient policing has far greater deterrent effects than increasing the use of incarceration, or indeed other sanctions

Recent research has demonstrated, however, that it is not always necessary to impose the most severe sanctions in order to maximise deterrent effects. Weisburd and Waring (2001), for example, found that financial penalties deter future offending by white collar criminals far more than does imprisonment. The process of detection, investigation and arraignment for a white collar offender is likely to produce similar deterrent effects as is actually serving a term of imprisonment.

Arguably, more imaginative sanctions ought to be applied to financial offenders. Braithwaite (1992, p. 170), for example, describes the utility of using so-called 'equity fines' in which companies are ordered to issue a certain proportion of new shares which are given to victims or to the state. For example, if a court ordered a corporate offender to is sue one new share for every 100 already issued, the market value of all shareholdings would be reduced by 1 per cent. The company would still be able to operate although shareholders would be penalised. Other possible sanctions include corporate probation, adverse publicity and community service. These are all able to be used within the existing sanctioning structure, although require a little imagination by prosecutors and judges.

4.0 CONCLUSION

This unit has identified some of the problems associated with using the criminal justice system to deter cross-border plastic card fraud. Some problems are capable of resolution through continued international cooperative efforts of law makers, police and investigators. Others, such as the eradication of jurisdictional safe havens will take longer to achieve and will require concerted international effort.

In the short-term, it remains necessary for cases of serious fraud to be investigated and prosecuted and for successful judicial outcomes to be widely publicised. Only then will that small group of organised criminals who make decisions about where, when and how to offend on the basis of some rational calculation, gradually come to realise that the return on their investment in perpetrating this type of crime may not be as great as they once believed it to be. Sanctions should, however, be applied appropriately. Often substantial terms of imprisonment may not be the most effective and efficient means of achieving deterrence.

Applying alternative sanctions, including those that entail civil and financial consequences, may be a better way of reducing plastic card fraud than always seeking to impose the most severe penalty of incarceration.

In addition, rather than increase spending on the wider use of incarceration, prospective offenders would more effectively be deterred through increased efforts at fraud prevention and enhanced rates of detection and reporting of offences. Increased spending on law enforcement activities might first be directed at detecting crime quickly and with certainty, and publicizing this fact. In addition, spending on education, training in business ethics, and fraud prevention initiatives would produce benefits in terms of reducing fraud to complement expenditure on the incarceration of individuals.

Plastic card fraud, and particularly those forms that involve an international component, need to be addressed using a range of strategies that extend from environmental measures that make crimes more difficult to commit to the use of judicial sanctions that seek to reduce the rewards derived from criminal conduct and to persuade those offenders who do make rational choices when deciding to offend, that the costs involved might, in fact, outweigh any benefits to be derived.

5.0 SUMMARY

- Plastic card fraud in the twenty-first century is no longer a domestic problem for financial institutions, merchants, consumers and law enforcement agencies but is now of global concern.
- The difficulties associated with prosecuting cross-border plastic card fraud are such that hard decisions have to be made about whether or not to embark on a prosecution in the first place.
- One consideration that rational-choice offenders may take into account in deciding whether or not to offend is the likelihood that they will be prosecuted. In the case of cross-border crimes where the offender is located in a different jurisdiction from the victim, a rational-choice offender may select a jurisdiction that will minimise the chances of prosecution.
- Turning to the question of harmonisation of laws, we have also seen some important developments take place in recent years that are relevant to the prosecution of financial crimes. International conventions need to deal not only with substantive laws relating to crimes of dishonesty, but also jurisdictional and procedural laws concerning mutual legal assistance.
- In Australia, the judicial punishments currently available in most jurisdictions are fines; restitution and compensation orders; forfeiture and disqualification (confiscation); unsupervised release (suspended, deferred, conditional sentences); supervised release (probation,

community service, intensive corrections) and custodial orders (either full time or periodic).

- Research supports the view that it is not the type of sentence which determines an offender's future criminal career, but rather various social and personal factors including access to employment and family and community support

6.0 TUTOR-MARKED ASSIGNMENT

Briefly discuss Article 6 of the International Convention on misuse of devices.

7.0 REFERENCES/FURTHER READINGS

Amnesty International (1998). People's Republic of China: The Death Penalty in 1997, Amnesty International, New York. /ASA/31702898.htm (visited 19-4-02).

Associated Press, (2000). 'Saudi Arabia Beheads Seven Nigerians'. http://cnn.com/2000/WORLD/meast/05/13/saudi.beheading.ap/index.html (visited 16 May 2000).

Ayres, I. & Braithwaite, J. (1992). Responsive Regulation: *Transcending the Deregulation Debate. New York: Oxford* University Press.

Braithwaite, J. (1992). 'Penalties for White-Collar Crime', in Complex Commercial Fraud.

Grabosky, P. N. (ed.), AIC Conference Proceedings No. 10, Australian Institute of Criminology, Canberra, pp. 167-71.

Braithwaite, J. and Drahos, P. (2000). Global Business Regulation,. Cambridge: Cambridge University Press.

Chan, J. (1995). 'The Limits of Incapacitation as a Crime Control Strategy', Contemporary Issues in Crime and Justice, New South Wales Bureau of Crime Statistics and Research, no. 25.

Clarke, R. V. and Felson, M. (1993). 'Introduction: Criminology, Routine Activity, and Rational Choice', in Clarke, R. V. and Felson, M. 9eds.), Routine Activity and Rational Choice', Advances in Criminological Theory, vol. 5, Transaction Publishers, New Brunswick, pp. 1-14. Council of Europe (2002), Convention on Cybercrime, European Treaty Series No 185, Budapest, 23 November 2001, Council of Europe, Starsbourg

Cornish, D. B. and Clarke, R. V. (eds.) (1986). The Reasoning Criminal: Rational Choice Perspectives on Offending, Springer-Verlag, New York.

Fox, R. and Freiberg, A. (1999). Sentencing: State and Federal Law in Victoria, 2nd ed., Melbourne: Oxford University Press.

Geurts, J. (2000). 'The Role of the Australian Federal Police in the *Investigation of High-Tech Crimes', Platypus Magazine: The* Journal of the Australian Federal Police, March, http://www.afp.gov.au/publica/platypus/mar00/intfrd.htm (visited 5 February 2001).

Ghosh, A. (2002). 'The Cybercrime Act 2001: Implementing the *European Union's Cybercrime Convention', Paper presented at* the RSA Conference, San Jose, 16-22 February.

Harding, R. W. (1990). 'Rational-Choice Gun Use in Armed Robbery: The Likely Deterrent Effect on Gun Use of Mandatory Additional Imprisonment', Criminal Law Forum, vol. 1, no. 3, pp. 427-50.

KPMG (2002). Fraud Survey, KPMG Sydney.

Krambia-Kapardis, M. (2001). Enhancing the Auditor's Fraud *Detection Ability: An Interdisciplinary Approach, Peter Lang,* Frankfurt am Main.

Lanham, D. (1997). Cross-border Criminal Law. Sydney: John Libbey & Co.

Lakeman, P. (2001). 'Mechanisms for International Cooperation: *Interpol's Universal Classification System for Counterfeit Payments Cards', in Broadhurst, R. G. (ed.), Proceedings of the* Asia Cyber Crime Summit, Hong Kong, 25 -26 April, Centre for Criminology, University of Hong Kong.

Levi, M. (2000). 'The Prevention of Plastic and Cheque Fraud', A Briefing Paper Prepared for the Home Office Research. London: Development and Statistics Directorate.

Lisker, J. S. (2001). 'Electronic Commerce Fraud: Risk Assessment and *Prevention', in Broadhurst, R. G. (ed.), Proceedings of the Asia* Cyber Crime Summit, Hong Kong, 25 -26 April, Centre for Criminology, University of Hong Kong.

McConnell International (2000). 'Cybercrime and Punishment? Archaic *Laws Threaten Global Information'*. http://mcconnellinternational.com/services/CyberCrime.htm (visited 30 January 2001).

Newton, J. (1995). Organised Plastic Counterfeiting, HMSO, London.
Page, F. 1997, 'Defining Fraud: An Argument in Favour of a *General Offence of Fraud', Journal of Financial Crime, vol. 4,* no. 4, pp. 287-308.

Sussmann, M. A. (1999). 'The Critical Challenges from International *High-Tech and Computer-Related Crime at the Millennium',* Duke Journal of Comparative and International Law, vol. 9, no. 2, pp. 451-89.

Tunnell, K. D. (1996). 'Let's Do It: Deciding to Commit Crime', in Conklin, J. E. (ed.), New Perspectives in Criminology, pp. 246-58, Boston: Allyn and Bacon.

United Nations (2000). Convention against Transnational Organised Crime, Document: Victoria, Attorney-General's Department 1988, Report of the Victorian Sentencing Committee, Melbourne: Government Printer.

Weisburd, D. and Waring, E. (2001). White Collar Crime and Criminal *Careers. New York: Cambridge University Press.*

Young, D. (1986). Translation of Beccaria's On Crime and *Punishments. Indianapolis: Hackett Publishing Co.*

Zimring, F. E. and Hawkins, G. J. (1973). Deterrence: The Legal Threat in Crime Control. Chicago: University of Chicago Press.