Electronic Security:
Risk Mitigation in Financial Transactions

Public Policy Issues

Thomas Glaessner, Tom Kellermann, Valerie McNevin*

June 2002
The World Bank

* Thomas Glaessner is Lead Financial Economist in the Financial Sector Operations and Policy Department of the World Bank, Tom Kellermann is a Data Risk Management Specialist, and Valerie McNevin is Security Information Officer and the Privacy Officer for the State of Colorado.

Contents

Abstract

This paper builds on a previous series of papers (see Claessens, Glaessner, and Klingebiel, 2001, 2002) that identified electronic security as a key component to the delivery of e-finance benefits. This paper and its technical annexes identify and discuss seven key pillars necessary to the fostering of a secure electronic environment. Hence, it is intended for those formulating broad policies in the area of electronic security and those working with financial services providers (e.g., executives and management). The detailed annexes of this paper are especially relevant for chief information and security officers responsible for establishing layered security.

First, the paper provides definitions of electronic finance and electronic security and explains why these issues deserve attention. Next, it presents a picture of the burgeoning global electronic security industry. Then, it develops a risk-management framework for understanding the trade-offs and risks inherent in the electronic security infrastructure. It also provides examples of trade-offs that may arise with respect to technological innovation, privacy, quality of service, and security in the design of an electronic security policy framework. Finally, it outlines issues in seven interrelated areas that often need attention in the building of an adequate electronic security infrastructure. These are (i) the legal framework and enforcement; (ii) electronic security of payment systems; (iii) supervision and prevention challenges; (iv) the role of private insurance as an essential monitoring mechanism; (v) certification, standards, and the roles of the public and private sectors; (vi) improving the accuracy of information about electronic security incidents and creating better arrangements for sharing this information; and (vii) improving overall education about these issues as a key to enhancing prevention.

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and should not be attributed in any manner to the World Bank, its affiliated organizations, members of its Board of Executive Directors, or the countries they represent.

Acknowledgments

The authors would like to thank the following individuals who have shared their time and background material. In addition, we thank them for their valuable written and oral input. They are Forrester Allison, Chris Bateman, Tim Burke, John Carlson, Joseph Cooper, Tony Chew, Dr. Dorothy Denning, John Farber, Rick Fleming, Douglas Foster, Karen Furst, John Frazzini, Edward Gilbride, Joi Grieg, Hugh Kelly, James H. Lau, Demetrios Lazarikos, Stephanie Lanz, Toby Levin, Peter MacDoran, Simon Martinez, Linda McCarthy, Sallie McDonald, Joe McLeod, Raj Nanavati, Kevin Nixon, Brian Palma, Dr. Joseph Pelton, Peter Penfiel, Bill Rogers, Ty Sagalow, James Savage, Troy Schumaker, Keith Schwalm, Leonardo Scudere, Don Skillman, Kurt Suhs, Gary Sullivan, Orson Swindle, Cornelius Tate, Dave Thomas, Mike Voorhees, Bob Weaver, Bob Wice, Bill Worley, and Paul Zanker. In addition to these individuals, many private organizations and public agencies took time to share their ideas with the authors, both in person and in a Global Dialogue held via the World Bank video conference facilities that included a discussion of e-security issues with officials from 20 countries in Latin America, Asia, and Africa.

Executive Summary

A previous series of papers on e-finance (see Claessens, Glaessner, and Klingebiel, 2001, 2002) identified electronic security as crucial to enabling electronic finance to meet business and consumer expectations and deliver the benefits provided by technology and leapfrogging. Even before the events of September 11, electronic security was a growing risk-management concern of banks and other financial services providers. But post-September realities require an altered approach to wireless technology. Its rapid growth in many emerging markets and its increased use, either with the Internet or on a freestanding basis, now demand a harder look at electronic security in the financial sector.

E-security touches the very heart of the new economy. For the first time since World War II, global markets and the global community can promise significant benefits to all. But the process of building a global economy demands discussion of important issues, such as how to define and protect privacy, what trust and confidence will mean and how to measure them, and how to determine security.

Due to the ever changing nature of technology, this paper does not treat all these issues nor does it attempt to provide definitive answers. Rather, it offers a view of what has transpired to date, the gaps that are opening in the electronic security area, and some possible approaches for bridging those gaps. It also acknowledges all the people around the globe who have worked diligently to resolve these and other issues in a fast-paced, constantly changing environment.

This paper and its technical annexes identify, define, and discuss seven key pillars needed to develop policies, processes, and an overall infrastructure that fosters a secure electronic environment for the financial services sector. It is intended for policymakers working with financial services providers, especially executives and security officers. The technical annexes reflect the views of many persons who are active in the electronic security industry; they should be of special use to those who administer electronic security systems, bank examiners who evaluate the adequacy of electronic security, and those who deal with the associated day-to-day risks inherent in electronic transactions.

What is electronic security?

Broadly speaking, electronic security is any tool, technique, or process used to protect a system's information assets. Electronic security enhances or adds value to a naked network and is composed of soft and hard infrastructure. The soft infrastructure components are the policies, processes, protocols, and guidelines that protect the system and the data from compromise. The hard infrastructure consists of hardware and software needed to protect the system and data from threats to security from inside or outside the organization. The degree of electronic security used for any activity should be proportional to the activity's underlying value. Security is a risk-management, or risk-mitigation, tool. Appropriate security means that the risk has been mitigated for the underlying transaction in proportion to its value. Given that the Internet is a broadcasting medium, constraints have to be added to target only intended recipients. As a result, the need for security is a constant of doing business over the Internet.

Electronic security will require more attention as new technology creates new risks and as technologies converge.

For purposes of this paper, e-finance is the use of electronic means to exchange information, transfer signs and representations of value, and execute transactions in a commercial environment. E-finance comprises *four primary channels*: electronic funds transfers (EFTs), electronic data interchange (EDI), electronic benefits transfers (EBTs), and electronic trade confirmations (ETCs).

Although e-finance offers developing market economies an opportunity to leapfrog, it is not without potential risks. Most of the crimes that take place over the Internet are not new—fraud, theft, impersonation, denial of service, and related extortion demands have plagued the financial services industry for years. But technology opens up new dimensions of depth, scope, and timing, enabling perpetrators to engineer with flexibility and specificity much greater disruption or theft than ever before. Technology creates the possibility for crimes of great magnitude and complexity to be committed very quickly. In the past, stealing 50,000 credit card numbers would have taken months, perhaps years, for highly organized criminals. Today one criminal using tools freely available on the Web can hack into a database and steal that number of identities in seconds. Or a perpetrator can steal a laptop containing a database of 400,000 names and their associated credit card information. These are the reasons e-security must be taken very seriously.

Because electronic data is invisible to the naked eye and because most people do not have computer skills, they may erroneously believe that stored information cannot be captured easily. In truth, it often takes few skills to access the data and manipulate, pollute, or steal it. Ironically, it may require even fewer resources and skills to protect the data in the first place. All four channels of e-finance are susceptible to fraud, theft, embezzlement, pilfering, and extortion. Equally important is that in many breaches the breakdown results from a failure to implement appropriate risk-management processes or from the use of off-the-shelf commercial software

Recent surveys suggest that in the United States, 57 percent of all hack attacks were initiated in the financial sector last year. The threat goes beyond financial and reputation loss to a potential backlash against electronic transactions and the increasing lack of trust that could occur among consumers. Identity theft is already the number one crime in the United States, and reported incidents of it are projected to more than double, from 700,000 last year to 1.7 million in 2005. Moreover, the cost to U.S. financial institutions is expected to increase 30 percent each year, to more than $8 billion in 2005. These problems have not been unique to free-standing financial institutions. For example, inadequate audits and underwriting processes within the U.S. system for electronic retail distribution of securities (Treasury Direct) nearly allowed one individual to compromise the whole system.

The new network-mediated economy paradoxically presents unparalleled opportunities for the creation of good outcomes or the perpetration of bad ones. Most countries are experiencing the dichotomies of the new economy. In assessing its promises and weighing these against potential pitfalls, policy and decision makers should appreciate that e-security is a prerequisite to transacting secure business on the Internet.

The electronic security industry is growing—becoming global—and will present public policy challenges even in the areas of competition policy, potential conflicts of interest, and certification.

For the most part, today's emerging global e-security industry did not exist before the early 1990s. It has grown in an attempt to retrofit the new defense-initiated technologies for unintended uses, that is, from using personal computers (PCs) and the Internet for finance and trade.

E-security companies and vendors generally fall into three categories: access, use, and assessment. Today's industry includes companies that provide active content monitoring and data filtering, develop intrusion detection services, place firewalls, conduct penetration tests to expose hardware or software vulnerabilities, offer encryption software or services, and create authentication software or services that use passwords, tokens, keys, and biometrics to verify the identity of the parties or the integrity of the data.

In addition to e-security, many vendors supply a multitude of interlinking services to the e-finance providers in many countries. These services include hosting companies, Internet Service Providers (ISPs), and providers of financial services. Telecommunication companies in many emerging markets are also often the key providers of cellular, satellite, and microwave services. These companies often have a stranglehold on access to telecommunication delivery channels, and because of the scarce skilled human capital, these companies of necessity often supply hosting services and de facto money transmission services. Just as important, they often provide certain electronic security services.

The cross-linking ownership of the e-security and e-finance industries raises many complex questions, such as the need to review competition policy as well as the potential for and ramifications of multiple conflicts of interest. In the case of competition policy, do the multiple roles played by telecom companies act to inhibit competition, particularly in emerging markets where the expertise to provide such services often resides in these companies? More important may be issues of conflict and integrity of the services provided as well as incentives to report security breaches accurately. For example, will a telecommunication company that provides hosting, Internet service, and e-security to a bank act on its own volition, with no regulatory mandates, to institute adequate electronic security measures or report intrusions accurately and in a timely fashion? Will such an entity be able to provide proper certification of digital signatures when it has business interests in so many conflicting areas? Moreover, such an industry structure with an extensive use of outsourcing will need to review the extent of downstream liability required by this complex set of vendors as the extent of liability can at least mitigate some of the incentives that can exist for important conflicts. In many countries, liability stops with the user—in this case, the financial services provider. Typically, contracts between financial entities and their providers use service-level percentages as a performance guarantee on a sliding-cost scale, but they do not build in sufficient remedies to address product performance from a security perspective.

The public interest case for regulation of the electronic security industry must be recognized. Important trade-offs exist between electronic security and such areas as costs, quality of service, technological innovation, and privacy. Formulation of regulation and policy needs to take explicit account of these trade-offs.

Traditionally, the telecommunications industry has been regulated as being essential to public health, interest, and welfare. Hence, a core component of its regulatory model was to expand service to give everyone access. In many countries, access to basic service is now considered a necessity of modern life. Historically, the financial services industry has been regulated by the premise that trust and confidence are paramount to the orderly movement of trade, goods, and money. And, given that a special trust is conferred on financial entities, they must conduct their business in a safe, sound, and prudent manner. Convergence of the telecommunications industry and the financial services sector through the Internet heightens the importance of and the necessity for sound public policy and informed regulation to ensure that government, business, and people continue to have access to secure financial services.

Beyond the issues raised by cross-linked ownership of the e-security and e-finance industries, there are even more basic issues to address in the design of a revised e-security public policy framework.

First, given the risks electronic security breaches pose to critical infrastructure, e-security is important in promoting and protecting public health and welfare. For example, the electronic economy is exposed to and dependent on the Internet and the public-switched network as its main transmission vehicles. In a related way, the critical elements of the electronic economy are integrally connected, from financial services to electricity through the phone system. Breaches can quickly disarm or even compromise such key infrastructure areas as telephones or electricity and detrimentally affect the payment system. September 11 provides an example, though less well-known hacker attacks have resulted in the loss of ATM or other financial services. Hence, in addition to the concerns raised by the structure of the growing e-security industry and the dependence of the e-finance industry on the continuous access to telecommunications. There is a fundamental public interest case for a government to regulate the e-finance industry and to ensure that it operates with at least a minimum of layered e-security.

Second, a market failure is occurring because inadequate incentives exist within the workplace, as well as within the regulatory and enforcement arenas, to require the timely and accurate reporting of electronic security breaches. Often, financial entities and corporations do not report losses, suspected losses, or breaches for fear of losing consumer or shareholder confidence. A recent survey suggests that institutions fail to report for the following reasons: they fear negative publicity; they want to ensure that their competitors will not use negative information; their managers are unaware that they can or should report events; they decide a civil remedy is best; or their information technology (IT) personnel are worried about their job security after an incident. Clearly, regulators have the upper hand in this dilemma. By requiring timely and accurate reporting with sufficiently strong penalties for failing to report, management and/or employees can be incentivated to report a breach incident to appropriate authorities.

Third, formulation of policy in this area must balance a number of complex competing concerns; in the end, electronic security cannot be seen as an end in itself but rather as only one aspect of risk management. Trade-offs exist between the costs of financial services provided, the size of transactions, and the sophistication of the electronic security arrangements that might have to be in place. Similarly, the quality of a financial service can suffer if security arrangements slow down transaction processing speed or result in other encumbrances for consumers of financial services. Technological innovation may also require that electronic-security-related regulations or legal statutes be as technology-neutral as possible. An example is digital signatures, often defined even in statutes as “requiring public key infrastructure”(PKI) when other authentication technologies might also be appropriate and should not be precluded by legislation or regulation. Finally, it is necessary to carefully weigh essential trade-offs between security as a protection component and the privacy element of access.

Any attempt to develop public policy to improve or establish electronic security needs to be built on at least the following seven important pillars: (i) an adequate legal and enforcement framework—which is not present today in many emerging markets; (ii) adequate arrangements to ensure electronic security of payment systems; (iii) an adequate supervision and prevention regime that creates better incentives to implement appropriate layered risk-management systems, including electronic security for financial services providers; (iv) encourage and promote a framework within which private insurance companies can insure against and monitor e-risk, thereby helping to improve standards in this area via the underwriting covenants they require; (v) develop certification standards and processes established with respect to digital signatures and, more broadly, to vendors operating in the electronic security industry; (vi) actions to improve the accuracy of information available about e-security incidents and the roles of the public and private sectors in this process; and (vii)educate citizens, employees, and management on security issues as a means of preventing e-security incidents.

Pillar I: Legal Framework and Enforcement

Countries adopting electronic banking or electronic delivery of other financial services (e.g., distribution and trading of securities) must incorporate electronic security concerns in their policies, laws, and practices, thereby allowing these to support secure operation of their institutions and to combat crime and cyber terrorism.

At a minimum, an e-finance legal framework should consist of the following:

- Electronic transactions law and electronic commerce law
- Payment systems security law
- Privacy law
- Cyber crime law
- Anti-money laundering law
- Enforcement infrastructure

Together, these six areas of law and enforcement address the ***basic relationships*** among all participants and the ***transactional activity*** that flows through the payments system. The cornerstone of an e-finance legal framework is to recognize the legal validity of consumer electronic signatures, transactions, or records. The legal framework should prefer technology-neutral solutions, provide basic consumer protections for electronically based transactional activity, promote interoperability, and address records retention.

Electronic Transactions Law. Discussed in greater detail later in the paper, electronic transactions law needs to define what is meant by an electronic signature, record, or transaction. It also needs to recognize the legal validity of each of these. It should be especially careful in defining an electronic signature. Definitions should apply to all non-consumer-related transactions and records. In the case of the electronic signature, the definition must be technology-neutral to the greatest degree possible to allow various technologies to provide solutions. The definition should further address the issue of record retention, because a valid signature is no good unless a valid retention process can support its validity when it is questioned.

Payment Systems Security Law. These statutes should identify, license, and regulate any payment system entities that directly affect the system, such as money transmitters and ISPs. They should provide that all such entities must operate in a secure manner so as to protect the integrity and reliability of the system. Further, they should require timely and accurate reporting on all electronic-related money losses or suspected losses and intrusions. And finally, they should require that the financial institution and related providers have sufficient risk protection. At minimum, they should encourage a shared-risk approach. Most countries have laws in place that regulate different components of the payments system. To date, however, no country has addressed comprehensively the electronic security challenges raised by payment systems.

Privacy Law. Privacy law should encompass data collection and use, consumer protection and business requirements, and notices about policy on information use. The European Union (EU) continues to be the leader in providing privacy protection to its citizens with the 1990 EU Directive on Data Collection. At a minimum, the privacy law should embrace the fair information practice principles of notice, choice, access, and security.

Cyber Crime Law. These laws should address abuses of a computer or network that result in loss or destruction to the computer or network, as well as associated losses. They should also provide the tools and resources needed to investigate, prosecute, and punish perpetrators of cyber crimes and, where needed, address the subject of adequate record retention to allow for electronic forensics and investigation. It is not possible to overstate the importance of such cyber crime legislation.

Anti–Money Laundering Laws. These statutes should define money laundering and require international cooperation in the investigation, prosecution, and punishment of such crimes pursuant to the guidance provided by the Financial Action Task Force (FATF). The FATF recognized the link between cyber vulnerabilities and money laundering when it modified its recommendations in 1996 to state, "Countries should pay special attention to money laundering threats inherent in new or developing technologies that might favor anonymity, and take measures, if needed, to prevent their use in money laundering schemes."

Enforcement. Perhaps more important than the legal framework will be the need to enforce the provisions of e-security laws within and across national boundaries. The fact that so many different types of computer or system related intrusions actually originate through activities conducted in countries with weak legal and enforcement regimes for electronic security, makes it essential that a broad international approach that relies on more homogeneous laws and enforcement actions across countries be put in place. More specifically a specific set of steps is needed. First, in many countries (both developed and developing), the civil and criminal penalties for unauthorized access to or tampering with computer systems are very lax and need to be significantly increased vis-à-vis their present status. Second, to more effectively prevent regulatory arbitrage and the use of countries with lax legal and enforcement infrastructure as staging grounds for undertaking cross border intrusions countries must, harmonize their approaches to cyber crime generally and to security-related crimes in particular. Third, access, availability, and interoperability should be the mantra for supervision and should guide enforcement efforts. Hence, the traditional regulatory structure must expand to include all entities that assist the financial institution throughout its information asset management cycle. This includes everything from ISPs to application service providers, software, hardware, monitoring detection, and assessment providers. Harmonization of penalties and sharing of information in enforcement are critical and require an exceptional international cooperative effort, given the active growth of organized criminal hacking syndicates that locate in jurisdictions with weak statutes and enforcement capacity.

Pillar II: Electronic Security of Payment Systems

Payment systems are among the most important components of any financial system. The operative questions for this paper are whether money transmitters or ISPs add risk to the payment system, and, if so, how that risk might be best mitigated. Any answers to these questions must address at a minimum the following five problems:

1. Lack of definition for money transmitters.
2. Lack of reporting requirements.
3. Limited or no regulation.
4. Limited or no warranties, indemnification, and liabilities.
5. Lack of security requirements for service providers.

Definition of a Money Transmitter. Generally speaking, a money transmitter is any commercial enterprise engaged in the transfer and exchange of monetary instruments and currency. Often these non-depository entities are involved in the "money service business" and serve as third-party automated clearinghouse providers.[1] The failure to regulate money transmitters is a symptom of the greater failure to define a new paradigm for money movement in

[1] These services may include money order issuance, wire transfers, currency exchange, and so on.

a sophisticated IT environment. Today, money or payment orders run through various channels, ranging from bits and bytes to snail mail, resulting in the continuing disintermediation of traditional banking. The phenomenal amount of money that flows *around* banks instead of through them has a significant impact on the global payment system, monetary policy, and economic forecasting. The failure to define money transmitters allows the methods of money movement to circumvent reporting requirements; run under the radar screen; and evade detection, regulation, enforcement, and punishment. Further, the failure of banks to report losses or suspected losses arising from IT use exacerbates these problems.

Reporting Requirements. Failure of regulators to require reports by these new money movers and to review and expand regulations to include new money movement vehicles permits unsafe and unsound activities in use of the payment system without check or prevention. Legislation should place an affirmative duty on executives[2] to report incidents, and the intentional failure to report should carry potential punishment.

Needed Regulation. Regulators should initiate supervision and enforcement over transmission vehicles. Financial service providers complain that this regulatory control will impede technological advances or undermine market innovation. But the primary reason cited by most people for refusing to use electronic transmission vehicles is fear that the information is not adequately protected. On the contrary, it will strengthen consumer confidence and market discipline. This, in turn, will pave the way for greater use of electronic financial systems. A simple licensing scheme for money transmitters, together with industry-sponsored certification criteria for financial services ISPs, is a time-tested approach to governing providers.

Indemnifications and Warranties. In addition to regulating these sectors, financial institutions should require warranties and indemnifications from businesses that create software and hardware or supply it to financial services providers or money transmitters. They also should require the companies that provide these products to be liable if losses occur as a result of software or hardware "holes." Off-the-shelf commercial software and hardware is widely known in the IT industry to be full of holes that can easily be used by hackers as "open doors." Any entity providing services or products to the financial services industry should either be held to a higher standard of care or required to disclaim up front that its product is not configured or otherwise appropriate for use in this sector. A variation on this solution is to require a disclaimer on hardware or software stating that it should not be used to create, move, or store confidential, privileged, or sensitive information and that if it is used for those purposes the manufacturer cannot be held liable.

The transference of risk and liability has been a hallmark activity of ISPs and host providers. Clearly, this is the norm for most countries, and in some countries it is the law. A clear change in business direction needs to occur in this area. Again, this change needs to be by government initiative, because it requires legal and regulatory intervention. In other words, regulators should require that a bank use an appropriately *secure* services provider, with the presumption that if it does not use such a vendor, it is operating in an unsafe and unsound manner. Initially, this may result in higher costs to financial services entities, but this expense would be more than offset by lower costs for insurance and ultimately by fewer losses and other costs of doing unguarded business over the Internet.

Standards for Service Providers. Service providers to the financial services industry also should be held to a higher standard than those not interacting directly with that industry. Again, this effort would go a long way toward building trust and confidence. For comparison, the United

[2] Particularly Chief Information Officers and Information Security Officers.

States has determined that "covered entities" providing services to the medical industry must meet certain standards. The Health Insurance Portability and Accountability Act (HIPAA) is a model piece of legislation governing electronic activity associated with health care and requiring that certain rules be promulgated to govern all aspects of the industry. At present, two of the eight rules are finished. These involve transaction codes, privacy, and security. HIPAA makes security a necessary prerequisite to providing services to the health industry, including the provision of any financial services. This model could be applied to the financial services industry.

Pillar III: Supervision and Prevention Challenges

Beyond the monitoring of the payments system and the related supervision of money transmitters is the need to revisit regulation, supervision, and prevention approaches to financial services providers that engage in electronic banking or provision of other financial services.

Capital Requirements. The new Basel guidelines for capital, especially those dealing with operational risk, do not address the problem of measuring either the risk to reputation or the strategic risk associated with electronic security breaches. Hence, there is a question of how best to measure a bank's operational risks when the information about computer security incidents is not accurate and when defining reputation damage is difficult, not to mention the needed adjustment to capital that would result from such a breach.

Given the problems involved in measuring capital adequacy in cases of electronic security risk, a more productive approach might be to use the examination process to identify and remedy electronic security breaches in coordination with better incentives for reporting such incidents.[3] In addition, authorities could encourage or even require financial services providers to insure against some aspects of e-risks (e.g., denial of service, identity theft) that are not taken into account within the existing capital adequacy framework. As the private insurance industry becomes more active in this field, this approach may be feasible, subject as well to the overall soundness and health of the insurance industry and its structure in emerging markets.[4]

Downstream Liability. The interlinked nature of financial services providers, money transmitters, and ISPs implies that the traditional regulatory structure must change or expand beyond its present configuration. This does not mean that a greater number of entities must come under the safety net, but rather that the legal or regulatory framework should create incentives for ISPs, hosting companies, application service providers, and software, hardware, and e-security providers to be accountable to the financial services industry. Liability must attach to these providers just as it does to the directors of those financial institutions that contract with them. Currently, these entities are not liable, and in some instances they are specifically excluded from liability under the law.

Supervision and Examination Processes. The Electronic Banking Group (EBG) was formed to make recommendations for needed additions, changes, or improvements in supervision and examination to accommodate the new technologies. The areas of supervision and examination will undergo major reorientations over the next few years. Just as the security industry experienced a paradigm shift with the mass introduction and dependence on PCs and the Internet, so must bank supervision realize that the center of gravity in the financial services industry is changing. Areas the EBG could evaluate further might include the following:[5]

[3] See the discussion of Pillar VI in this executive summary.

[4] In many emerging markets, the insurance industry itself may need to be restructured and be stable; however, cross-border provision of such coverage may be an option.

[5] The EBG and certain regulatory/supervisory agencies (OCC, MAS, FSA, HKMA) are already starting to take a more proactive approach.

- The means used to examine the IT systems of banks or other financial services providers in order to modernize the examination approach. It is important to determine whether the financial services provider and all related vendors use an adequate layered electronic security system (see Annex I).
- The institution's current documented security program. At a minimum, this requires that policies, practices, and processes be documented for risk assessment, monitoring, remediation, audit, and reporting.
- The procedures used to identify and assess entities that provide a data processing or money transmitter service to the institution. It is important to review documented internal criteria for reporting incidents to ensure the criteria are meaningful and sufficient for enforcement actions.

Coordination of agencies within and across borders. One important issue facing most countries is the need to improve the sharing of information across and among their regulatory and law enforcement agencies. Many countries have a number of entities for gathering critical information, but often it is not shared within a country or across nations (sometimes for legal reasons). Treatment of information-sharing between agencies in different countries or within a country is beyond the scope of this paper. But clearly, as a government tries to leverage its scarce resources in order to regulate and battle crime in this new environment, such warehoused information presents both an asset and a potential liability in what might be termed institutional "sharing wars." Improvement in this area will require joint enforcement actions and much greater cross-border cooperation.

Pillar IV: The Role of Private Insurance as a Complementary Monitoring System

Because financial supervisory agencies are still in the process of developing their regulatory standards, and because of the difficulties of monitoring complex transactions with rapidly changing technologies, it is important to seek complementary private solutions to the monitoring of risk. The insurance industry already is playing a role in this area despite the defects inherent in the underlying information used to price e-risk coverage. Within the next few years, in the United States market alone, the growth in e-commerce liability insurance and, specifically, e-risk coverage is likely to become quite large and may total as much as $2.5 billion annually.

The insurance industry is developing new products such as the above-mentioned e-commerce liability and e-risk coverage. Still in its early development, this insurance has problems in first- and third-party coverage. The pricing of cyber-risk insurance also is in need of fuller development, but to accomplish this, the insurance industry requires a better base of information on security breaches and associated risk. In addition, current underwriting practices for this form of insurance have paid insufficient attention to the special risks that wireless technologies bring to the delivery of financial services (see Annex III). Insurance providers need to require that explicit electronic security standards for wireless technology be identified and used to mitigate these risks before they underwrite e-risk policies.

The global insurance industry can increasingly act as an important force for change in global electronic security requirements. First, it can strive to improve the minimum standards for electronic security in the financial services industry. The global insurance industry should strongly advocate the use of enhanced layered electronic security as a business prerequisite, as suggested in Annex I. Second, insurance companies can require that financial services entities use vendors that meet certified, industry-accepted standards to provide electronic security services as a way of mitigating their risks of underwriting coverage. Third, insurance companies should encourage regulators to require that financial services entities both provide information and improve the quality of data and information on incidents so they can better actuarially measure e-risks and return on investment. Finally, the industry should promote solutions that require e-security vendors and other e-enabling companies (hosting, etc.) to engage in risk sharing and in carrying appropriate liability.

Pillar V: Certification, Standards, and the Roles of the Public and Private Sectors

Both public and private entities must work cooperatively to develop standards and to harmonize certification and licensing schemes in order to mitigate risk. Two categories that require particular attention in terms of certification deal with electronic security service providers and transaction elements.

A necessary first step in securing e-finance is to require licensing by financial regulators of vendors that directly affect the payment system, such as money transmitters or ISPs. The next step is to require the industry to certify vendors that provide electronic security services. Many vendors already offer some type of certification, and recently the security industry has developed a Security Expert certification. Industry experts should review this certification process and, if they deem it appropriate, endorse it. By using certification, the industry benefits by providing the consuming public with a recognized structure, accountability between the industry and its self-proclaimed experts, and a means of separating the approved expert from the self-authorized expert. It also elevates the craft to a professional status and creates an incentive for the industry to both raise and protect standards. Certification is a time-tested approach from which numerous professional groups, such as accountants and lawyers, have benefited.

A second area to address is certification of such transaction elements as electronic signatures. The value certification brings to a transaction in part depends on who or what provides the certification and on the elements that are being certified. In general, certification would provide an enhancement of an existing service, such as that of notaries. Or it could be offered by a quasi-private entity, such as a post office, or a private entity, such as a bank. Each of these scenarios, however, presents unique structural and governance issues. For example, post office certification would require the use of registration agents and storage to maintain a large information repository. By contrast, in many countries private companies (financial services providers or nonfinancial companies) may be better equipped to provide the information infrastructure required to act as certification agents or to provide cross-certification checks for a fee. This, too, necessitates the creation of an internal governing structure and an appropriate repository and record-retention scheme to house and protect the digital information for perhaps decades.

The important element essential to any of these scenarios is that certification structures located in different jurisdictions must consistently provide the same attributes to the transaction and that a certifier’s scope of authority and liability must remain consistent across jurisdictional borders. Today, a significant lack of consistency exists in the means by which certification structures are organized under the law in the various states. For example, a certification authority in Utah has a very different scope of authority and operates under a different liability structure from a certification authority in Florida or Illinois. Thus, if a transaction is initiated in Illinois but concluded in Utah with a party in Florida, one should be acquainted with the laws of each state in order to appreciate the ramifications of using certification authorities in each state.

The consistency gap is even greater between different countries. At present, approximately 50 countries have some type of electronic commerce statute in place. Only a portion of these, however, recognize or address the issue of certification authorities, and the laws that exist contain no standardized or harmonized provisions as to the roles and responsibilities, much less the liabilities, of these authorities. In addition, record-retention requirements often are undefined. The irony is that an information economy cannot thrive without a living, accessible, and documented institutional memory.

As technologies continue to mold worldwide convergence and integration, authorities will be under increasing pressure to enter into special treaties and reciprocal agreements that allow for cross-certification between countries. Malaysia, Singapore, and Hong Kong have set up such arrangements. These may well serve as models for other countries.

Transactional certification is used primarily to ensure that the underlying transaction will be honored. The main barrier that continually haunts the transactional side of e-commerce is the issue of nonrepudiation. It is essential that the parties "know" each other and, more important, know that each will perform and that each has legal remedy if the other fails to do so. The value of e-commerce is significantly undercut if this function cannot be provided in a simple, cost-effective manner. This function is accomplished in the electronic world by verifying or authenticating a party's identity. Technology offers numerous ways to insert such functions into the e-process. Although the use of PKI technology and certification authorities is often touted as the only accepted means of ensuring security, it is necessary to consider also the costs, the cumbersome structure associated with PKI, and the legal inconsistencies associated with certification authorities. It is also important to analyze the benefits of using alternative solutions, such as biometrics or digital time stamps. The critical element is that the solution be consistent across borders in terms of scope and liability, no matter what technology is used to perform the function.

Pillar VI: Accuracy of Information on E-Security Incidents and Public-Private Sector Cooperation

The lack of accurate information on e-security incidents is the result of the lack of incentives to capture the data, measure it, and inform. At worst, the failure to inform is tantamount to a breach of ethics; at best, it is a failure to notice.

Electronic security would improve worldwide through the creation of a set of national and cross-border incentive arrangements to encourage financial services providers to share accurate information on actual denial-of-service intrusions, thefts, hacks, and so on. Today, ample evidence shows that no accurate base of information exists either within or across countries. Not only does this limit awareness but, even more important, it can limit the provision of private sector solutions (including insurance). This lack of information may even be acting to increase the cost to companies and financial services providers of insuring against such risks.

Greater public-private sector cooperation is needed in this area. Such examples as the Internet Security Alliance, the Forum of Incident and Response Security Teams (with 56 worldwide offices), the electronic crimes task forces in the United States, the InfraGard program run by FBI field offices, and the Computer Emergency Response Team (CERT) run by Carnegie Mellon have shown that cooperation results in greater information sharing among law enforcement and private providers of financial services. A common element in all these programs is a reliance on trust, in that the law enforcement or academic entities involved tell respondents they will not divulge their identity. Critical to any global solution will be for a universally trusted third party to administer a global base of information relating to electronic security incidents. In this area, the role of multilateral agencies to facilitate cooperation deserves examination. It is axiomatic that the more "connected" the economy becomes, the more important it is for each element to bear its portion of the burden. Today's financial services industry was founded as an integrated system. The technological changes of the past decade have expanded and heightened the interdependencies of that system.

Pillar VII: Education and Prevention of E-Security Incidents

Statistical analysis reveals that in many countries throughout the world, more than 50 percent of electronic security intrusions are still carried out by insiders. An uneducated or undereducated workforce is inherently more vulnerable to this type of incident or attack. By contrast, a well-educated workforce that is conscious of security issues can effectively add a layer of protection.

Educational initiatives will have to be targeted to financial services providers (both systems administrators and management), to various agencies involved in law enforcement and supervision, and to actual online users of financial services. Actions might include the following: improvement of awareness and education of financial sector participants about cyber ethics and appropriate user behavior on networked systems; creation of institution-wide e-security policies on appropriate behavior and the corresponding channels for reporting intrusions or incidents in close coordination with any effort to improve worldwide information on intrusions; development of awareness in the banking community in emerging markets about the need for "incident response plans" in case an incident transpires; facilitation of cooperation and transfer of know-how among law enforcement entities, financial intelligence units (FIUs), and supervisory agencies in developed and emerging markets via such devices as more active exchange programs between personnel; design of well-focused courses for examiners under the auspices of the Financial Stability Institute or other training centers; and development of a cross-border university outreach program to promote the training of future e-security professionals while also improving the education of online users of financial services.

Provisos

This paper treats a rapidly evolving area using a cross-disciplinary approach, integrating the economy, law, and technology as appropriate.

Because of its rapid growth, e-security is often wrapped in myth. Most countries, including those that have greater experience dealing with it, still know little, and emerging markets know even less. The paper focuses relatively more attention on lessons learned in the United States because it is considered the birthplace of the Internet and has had a longer time to experience its benefits and pitfalls, as well as to create some standards.[6] Just as important, the paper looks at the experiences and efforts of certain advanced economies in Europe, as well as of countries in Asia and South America. Clearly, however, much greater effort needs to be mounted to understand the specific problems of emerging markets in this area as well as to identify critical areas of legislation and relevant institutional arrangements needed to improve electronic security worldwide. Without such efforts, the great potential offered by adopting electronic finance and commerce can be significantly compromised, because the trust and confidence of market participants—so critical to transacting via the many different technologies now being created—will be detrimentally affected.

[6] Historically, the Internet was derived from ARPANET, which was designed in 1969 by the Advanced Research Projects Agency, Department of Defense.

I. Introduction

Is it a fact…that, by means of electricity, the world of matter has become a great nerve, vibrating thousands of miles in a breathless point of time? Rather, the globe is a vast head, a brain, instinct with intelligence! Or shall we say it is itself a thought, nothing but a thought….—Nathaniel Hawthorne, 1851.

Even before the events of September 11, electronic security was a growing area of concern for banks and other financial services providers in managing daily operational risk. Now, because of the rapid growth of wireless technology and its increasing use in providing financial services in emerging markets, either in coordination with the Internet or on a freestanding basis, there is even more demand for a careful look at issues related to electronic security.

This paper has three central objectives. The first is to define electronic security, discuss why this issue is becoming important worldwide, and characterize the players in the burgeoning worldwide electronic security industry. The second is to offer an economic incentive framework to use in addressing the problems posed by electronic security, with particular attention to financial services provided by banks. The third is to identify seven distinct pillars of reform that every country should construct and maintain to develop a secure electronic environment.

In meeting these objectives, the paper addresses the following public policy questions relevant to the future security of the global financial system:

- Are financial services providers given proper incentives to fully share timely and accurate information with law enforcement on security breaches? If not, is there a form of market failure taking place in this area within the financial services industry? What actions might be taken to facilitate public-private cooperation to remedy the situation? (See Sections II, III, and X.)
- What kinds of changes or additions to the legal and regulatory framework will be consistent with proper law enforcement within and across country boundaries? (See Sections V and VI and Annex IV.)
- What role should government play in setting policies, standards, and guidelines for e-security? How can it strike the proper balance between fostering technological innovation and establishing e-security standards? (See Sections IV and V and Annex IV.)
- What role should government play in regulating and supervising not only financial services providers but also third-party providers, such as money transmitters, hosting companies, ISP providers, and electronic security vendors? (See Sections VI and VII.)
- How should electronic records or transactions be verified or authenticated? What role should the government and the private sector play in certification? (See Section IX and Annex II.)
- What role can the private insurance industry play, especially in emerging markets, which often lack extensive human capital and capacity in regulatory agencies? Can it offer incentives to guide business toward a risk-management and risk-mitigation approach? How can layered security help in monitoring the operational and other risks created by electronic security breaches? (See Section VIII.)
- What roles can the government, private market participants, and the electronic security industry play in accurately measuring the extent of electronic security risk within and across countries? How can institutions improve their information and databases from which to measure this risk? (See Section X.)

- How can complementary and reinforcing actions be taken to ensure better electronic security in emerging market countries where regulatory, supervisory, and enforcement institutions are not strong? (See Sections IX, X, and XI.)

The answers to many of these questions are interrelated, and this paper approaches them in a systemic manner. The annexes offer a more detailed and technical analysis of the issues. Included also is a glossary of terms. Hence, the paper is intended for those formulating broad policies in the area of electronic security, those working with financial services providers (e.g., systems administrators in these entities), vendors of electronic security or other products (i.e., front-end Internet platforms provided by a hosting or portal company) that outsource to such financial services providers, and other participants in what is becoming a global electronic security industry.

The paper is divided into 11 sections. Each of sections II through XI addresses one set of the questions raised above. Section II defines electronic finance and security as used in the context of this paper; it explains why these issues will increase in importance as dependence on new technologies spreads into emerging markets and leapfrogging becomes a reality. Section III characterizes the functional categories of the global electronic security industry and describes its links to e-finance. Section IV delineates a risk-management framework for thinking about electronic security and outlines the elements necessary for policy development to ensure adequate electronic security. Section V outlines legal and enforcement issues. Section VI examines the complexities of electronic security with respect to payment systems and money transmitters. Section VII examines supervision and prevention of security breaches, including new approaches to oversight and inspection of security systems at financial services providers or nondepository institutions that act as money transmitters. Section VIII explores the opportunities for private insurance to participate in creating a risk-sharing approach to electronic security. Section IX examines certification issues within the electronic security industry, as well as the specific topic of electronic messages or signatures and the appropriate role of the government. Section X suggests possibilities for developing public-private partnership to improve the accuracy and availability of information about electronic security incidents. Section XI examines education as a key to improving protection against e-security incidents.

This paper treats the rapidly evolving are of electronic security from a perspective of technology. Too little is known about this subject in emerging markets. The paper focuses more attention on the United States, because the Internet originated there and because the defense and law enforcement agencies there have more experience in ensuring electronic security. It also focuses on some of the more advanced economies in Europe, as well as on Singapore and Hong Kong, to examine how electronic security issues have been addressed in those areas. Clearly, more research is needed to understand the specific problems of emerging markets as well as to identify critical areas of legislation and relevant institutional arrangements needed to improve electronic security standards worldwide. Unless it protects its information assets, the great potential electronic commerce offers can be significantly compromised.

II. What Is Electronic Security and Why Is It Needed?

Definitions of E-Finance and E-Security

To understand the need for electronic security, one must first precisely define what is meant by electronic finance. For purposes of this paper, e-finance is the use of electronic means to exchange information, to transfer signs and representations of value, and to execute transactions in a commercial environment. E-finance comprises four primary channels: electronic funds transfers (EFTs); electronic data interchange (EDI); electronic benefits transfers (EBTs); and electronic trade confirmations (ETCs).

EFT, which began in the early 1960s, is the oldest form of electronic money transmittal. The amount of money moving by EFT is $2 trillion per day and growing. The volume of EFT usage worldwide is 677,411,204 transactions.[7] The second oldest form of electronic money movement is EDI. EDI is used to effect money payment orders and bar coding. Bar coding is operational in more than 70 countries worldwide. Its use has doubled in the past five years and is equal to 50 to 75 percent of purchases worldwide. The third oldest channel is EBT. Benefits have been transferred electronically for a decade in more than 37 countries worldwide, including many emerging economies. In the United States alone, EBT moves $500 billion in cash entitlements, such as food stamps, Social Security payments, and child assistance benefits. The total volume of EBT transactions in the United States is 568,981,051 annually.[8]

E-security can be described on the one hand as those policies, guidelines, processes, and actions needed to enable electronic transactions to be carried out with a minimum risk of breach, intrusion, or theft. On the other hand, e-security is any tool, technique, or process used to protect a system's information assets. Information is a valuable strategic asset that must be managed and protected accordingly. The degree of e-security used for any activity should be proportional to the activity's underlying value. Thus, security is a risk-management or risk-mitigation tool, and appropriate security means mitigation of the risk for the underlying transaction in proportion to its value.

The need for security is a constant of doing business over the Internet because, in essence, the Internet is a broadcast medium. E-security enhances or adds value to a naked network and is composed of both a "soft" and a "hard" infrastructure. Soft infrastructure components are those policies, processes, protocols, and guidelines that create the protective environment to keep the system and the data from compromise. The hard infrastructure consists of the actual hardware and software needed to protect the system and its data from external and internal threats to security.

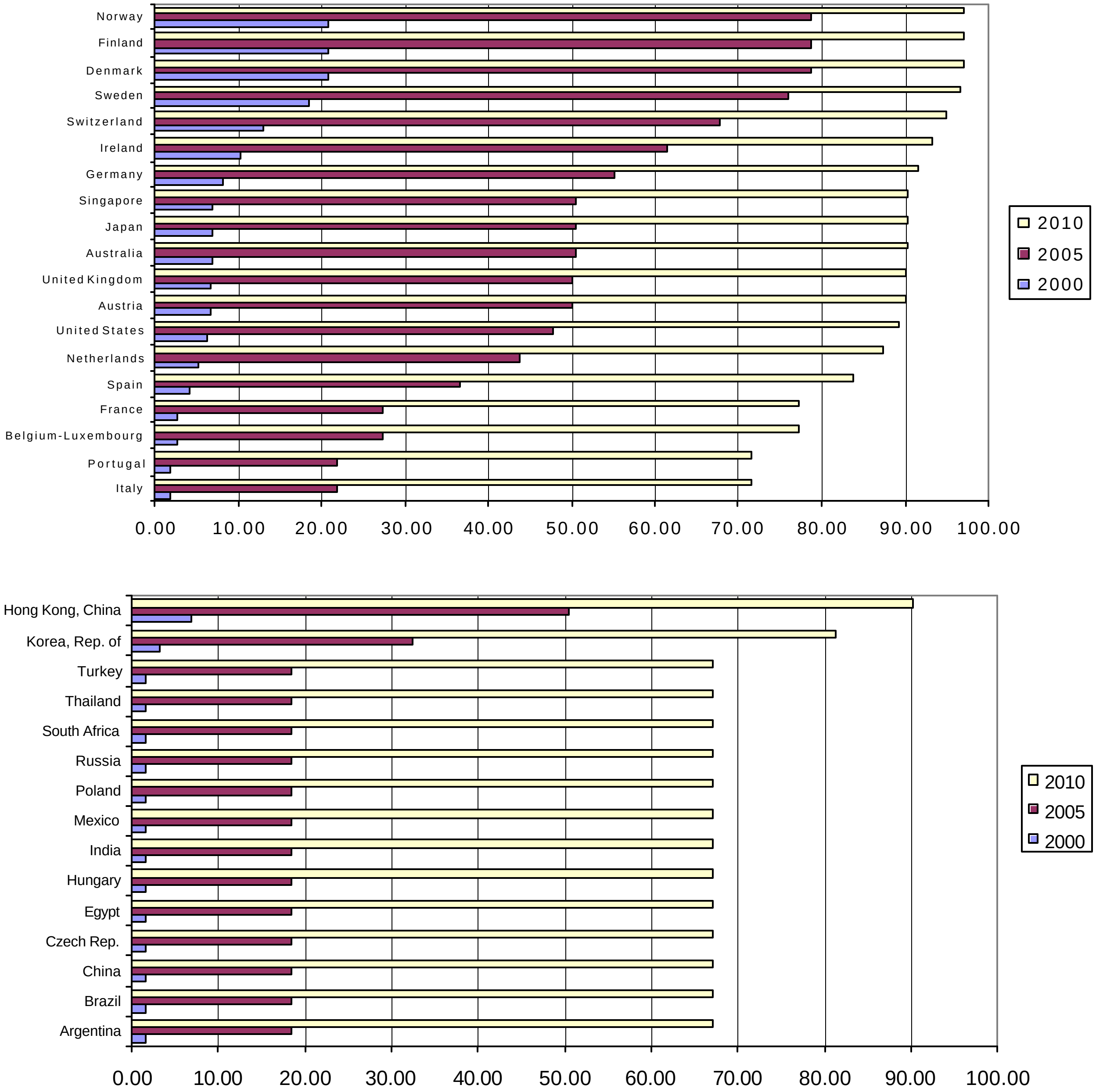
The Potential Growth of Electronic Transactions

The volume and variety of electronic financial services have increased significantly, and use of the electronic medium to do business, whether online or through remote mechanisms, has spread rapidly over the past decade. Countries, not just consumers, are increasingly getting connected. As is evident in Figure 1, "these new technologies not only allow countries to leapfrog in connectivity, they also open new channels for delivering e-financial services" (Claessens, Glaessner, and Klingebiel, 2001). Since the mid-1990s, investment in banking technology has focused on online banking and brokerage services to increase convenience and also to reduce costs.

[7] U.S. Department of the Treasury 2001.

[8] U.S. Department of the Treasury 2001 statistics.

Figure 1. E-Finance Penetration: 2000 and Projected Rates for 2005 and 2010



Note: The figures show projections based on takeoff years with better connectivity. The projections assume that all emerging markets have the same connectivity rating as in today's lowest-ranked industrial country, 6 (or better if their current rating is already higher); thus, the projections lead to the same minimum level of penetration in each emerging market.
Source: Authors' calculations.

Concurrent with these realities, four new technology-related financial services industry trends have occurred: outsourcing, open architecture, integrated strategies, and new methods of e-payment. The new trends have been driven by considerations of cost reduction and need for improvement in quality of service, yet in the process of putting them in place, security issues have too often been presumed to be less important or sometimes taken for granted. Figure 1 illustrates the projected rates of e-finance penetration worldwide. Fraud rates are more than 83 times higher[9] than those experienced by the bricks and mortar merchants. According to Meridien Research, online credit card fraud totaled $9 billion in 2001.

By 2005, the share of banking that is done online could rise from 8.5 percent to 50 percent in industrial countries, and from 1 percent to 10 percent in emerging markets. With better connectivity, online banking transactions in emerging markets could rise even further to 20 percent by 2005 (Glaessner, Claessens, and Klingebiel, 2001). Some estimate that $6.3 trillion of bank-to-bank transactions will be online by 2005.[10]

A parallel trend to the global use of e-finance is the adoption of new technologies that can act to expand the scope for electronic finance and access to financial services. Emerging markets increasingly find it more advantageous to use "new" technologies, such as wireless cellular technology, for e-finance as opposed to the Internet. Table 1 indicates that in a variety of emerging markets, wireless technology, as measured by cell phone penetration, is rapidly outstripping Internet penetration.

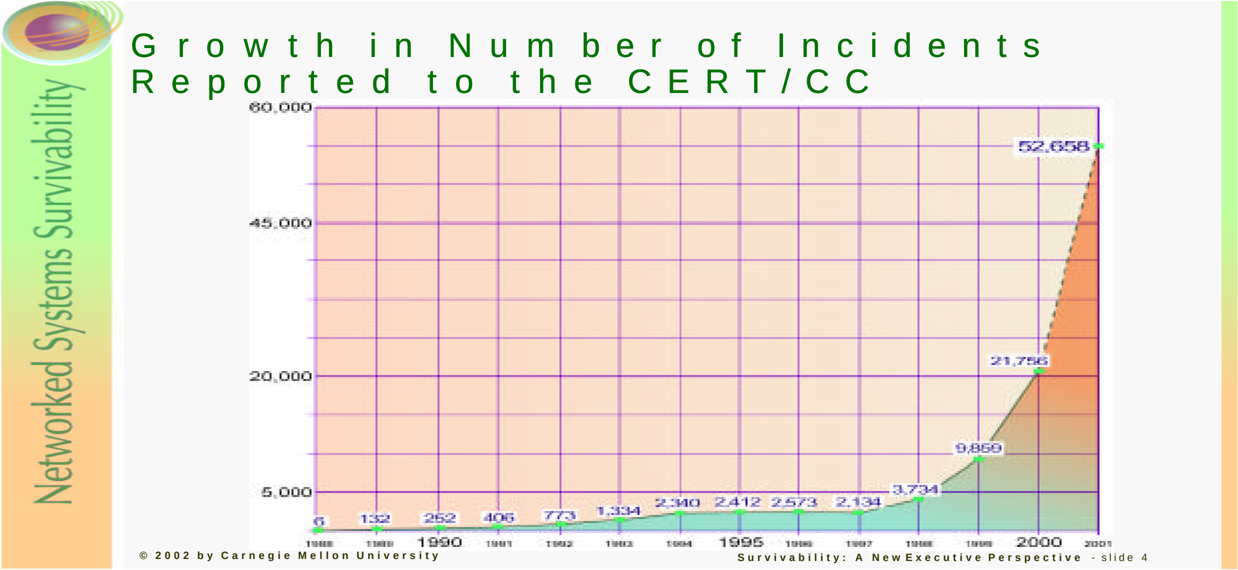
Table 1. Global Connectivity Trends

Country	*Number of mobile phone subscribers (Millions)*	*Percentage of population who are mobile or cellular subscribers*	*Percentage of population who are Internet users*
Developed Countries [a]	**30.0**	**56**	**32**
Australia	8.6	**45**	**35**
Finland	3.7	**72**	**38**
France	29.1	**49**	**14**
United States	109.0	**40**	**35**
United Kingdom	43.5	**73**	**30**
Developing Countries [a]	**6.9**	**7**	**2**
Brazil	23.2	**14**	**3**
Bulgaria	.6	**7**	**5**
Cambodia	.1	**1**	**<1**
China	84.5	**7**	**2**
Egypt	1.4	**2**	**1**
Guatemala	.7	**6**	**<1**
India	3.6	**< 1**	**<1**
Indonesia	3.7	**2**	**<1**
Mexico	14.1	**14**	**3**
Philippines	6.5	**8**	**3**
Republic of Korea	26.8	**57**	**40**
South Africa	8.3	**19**	**5**

Source: International Telecommunications Union, *World Telecommunications Indicators Database 2000.*
a. These are averages for developed and developing countries respectively.

[9] *The Myth of Online Payments*. Mike Voorhees (2002).
[10] Jupiter Communications (2001).

Figure 2. Increase in Incident Reports



Growth in Number of Incidents Reported to the CERT/CC

© 2002 by Carnegie Mellon University

Survivability: A New Executive Perspective - slide 4

The Risks of New Technologies

With the benefits of new technology also come new risks (see Figure 2). Table 2 shows that since 1995, incident reports increased 200 percent between 2000 and 2001 in the United States alone. Technology facilitates more efficient and quicker ways to commit old crimes such as fraud and theft. Remote access, high-quality graphic s and printing, and new multipurpose tools and platforms provide greater means to commit such crimes as theft and impersonation online.[11] Disturbingly, as the technology becomes more complex, a perpetrator needs fewer skills to commit these crimes. For example, the art of online penetrations (i.e., hacking) was once a highly sophisticated skill. The information age, however, has permitted a breeding ground for underground hacker Web sites that now supply dubious individuals with the multifaceted tools necessary to break into financial platforms. Such Web sites as www.astalavista.box.sk and www.attrition.org supply complex malicious codes and viruses that enable novice users to penetrate banking systems.

The most frightening aspect of the convergence of technology and crime is the speed and magnitude of the crimes that can be undertaken. For instance, in the past it would have taken months or perhaps even years for highly organized criminals to steal 50,000 credit card numbers. Today, one criminal using tools that are freely available on the Web can hack into a database and steal that number of identities in seconds. Or a perpetrator can steal a laptop containing a database of 400,000 names and their associated credit card information. These are the reasons e-security must be taken very seriously.

Although e-finance offers an opportunity for developing market economies to leapfrog, it is not a panacea. The Internet Data Corporation (www.idc.com) recently reported that more than 57 percent of all hack attacks last year were initiated in the financial sector.

[11] Ibid.

Traditional risks have thus been reshaped. In the physical environment, frauds traditionally were paper-based or people-based, whereas the following are the means most often used to commit crimes online:

- Message interception and alteration
- Unauthorized account access
- Identity theft[12]
- Manipulation of stocks and bonds
- Extortion
- Unauthorized system access (e.g., system damage, degradation, or denial of service)
- Industrial espionage
- Manipulation of e-payment systems
- Credit Card Theft[13]

The tremendous growth in open networks has created a penetrable electronic environment akin to a circle of Swiss cheese pieces. Financial institutions are increasingly relying on technology to process, store, and retrieve data, but advances in computer hardware, software, and communications technology increase the financial industry's vulnerability to internal and external attacks. Without strong security controls, banks risk the possibility of financial loss, legal liability, and reputation harm.

The insecurity of the Internet further exposes financial institutions to undetected, global, and virtually instantaneous attacks on internal systems and proprietary information. This includes attacks by foreign governments and terrorists, as well as attacks by criminals or hackers originating domestically. Banks and vendors with weak security controls are susceptible to business disruptions, theft of data, sabotage, corruption of key records, and fraud. The development of wireless Internet access will further compound the problem (see Annex III) by enabling foreign governments, terrorists, criminals, and hackers, singly or in concert, to operate in countries that do not have the advanced communications infrastructure or adequate security protocols in place. Hence, building awareness now of the criticality of the risks associated with e-finance and promoting industry use of aggressive mitigation is crucial.

Despite the relative lack of accurate information about actual intrusions and associated losses, Table 2 highlights some the most pervasive venues for electronic attacks in the area of e-financial services that have been publicly documented. The most frequent problems in this arena are (i) insider abuse, (ii) identity theft, (iii) fraud, and (iv) breaking and entering, often conducted by hackers.

[12] Refer to *ID Theft: When Bad Things Happen in Your Name*, an FTC publication for further information on how to mitigate the damage caused by the theft of one's personal information.

[13] Credit card theft is exploding via the Internet. Internet Chat Rooms facilitate the sale of stolen credit cards. Most hackers utilize online casinos to set up accounts, close them and them collect the illicit funds.

Table 2. Reported E-Security Intrusions

Date of Attack	*Compromised financial and e-commerce entities*	*Name of hacker, group, or malicious tool*	*Various losses sustained because of the intrusion into the financial entity's networks*
Sept. 18, 1995	Citibank[1]	Vladimir Levin	$ 10,000,000[2]
Mar. 1, 2000	U.K., U.S., Thailand, and Canada's e-finance and e-commerce sites	Alias "CURADOR"	28,000 accounts compromised, with total losses exceeding $3.5 million.[3]
Mar. 15, 2000	Internet Trading Technologies[4]	Abelkader Smires	Denial-of-service attacks that caused major disruption of trading on the NASDAQ.
Aug. 10, 2000	Bloomberg[5]	Oleg Zesev and Igor Yarimaka	Broke into the Bloomberg computer system in Manhattan in an attempt to extort $200,000.
Dec. 22, 2000	EggHead[6]	Eastern European groups	Hackers compromise database of thousands of credit cards; on Christmas Eve, many of the cards were then "salami sliced."[7]
2001	Hong Kong	Various Hackers	Eight cases of e-banking theft were recorded in the year involving the loss of over $4.4M. [8]
Mar. 8, 2001	40 domestic e-banking and e-commerce sites	Eastern European criminal syndicate	Intruders stole credit card account information and other data by exploiting a Windows NT security flaw; the National Infrastructure Protection Center labeled this attack the "largest Internet attack to date."[9]
Apr. 12, 2001	VISA	Eastern European groups	Intruders gained access to its computer network in the U.K. and later demanded ransom for data obtained in the virtual break-in; company received a ransom demand of £10 million.
Jun. 5, 2001	Central Texas Bank[10]	Vasilly Gorshov and Alexey Ivanov	They had access to the bank's system for six months before they were detected.
Jul. 6, 2001	S1 (a host company)[11]	Investigation ongoing	The compromise of more than 300 banks and credit unions whose systems were hosted by S1.[12]
Jul. 14, 2001	Australia's Online Trading Systems	Black Orifice—Trojan Horse	Account data of more than 40,000 of their clients was compromised.
Aug. 21, 2001	Riggs Bank, First Virginia Banks, SunTrust, and Visa	Investigation ongoing	The account information of more than 4,000 account holders from these banks who used Visa debit cards was compromised; banks were forced to cancel all debit cards. [13]
Sept. 20, 2001	Deutsche Bank[14]	Nimda worm	Unknown—costs of breaches indeterminable.
Feb. 7, 2002	U.S. Treasury Direct [15]	Louis Lebaga	$158 million—Lebaga was apprehended only after attempting to steal $1.3 billion more five days later.
Mar. 1, 2002	Prudential Insurance Company	Donald McNeese	McNeese was arrested for the theft and credit card scam stemming from the hack of Prudential's database, compromising 60,000 personal records of employees there. [16]
Apr. 5, 2002	State of California, Payroll database	Investigation Ongoing	The hacker copied 265,000 state employee account names and social security numbers, thus making them vulnerable to ID theft. [17]
Apr. 12, 2002	Republic Bank	Investigation Ongoing	The hacker copied 3,600 bank customer account names and files, thus making them vulnerable to ID theft; by exploiting S1's (the hosting company's) servers, he was able to compromise the accounts of these customers.[18]
June 19, 2002	Singaporean Bank DBS	A Chinese National[19]	The hacker siphoned over $31,000 from 21 various accounts, transferred $35,000 into his own account, withdrew the money at a bank branch and then fled the country. The whole operation took less than two hours.

Notes:

1. "Bank's Security Chains Rattled." *Financial Times*. Sept. 20, 1995. www.ft.com
2. Of the $10 million lost, all but $400,000 was recovered.
3. National Infrastructure Protection Center Major Investigations Web site: www.nipc.gov/investigations/curador.htm.
4. National Infrastructure Protection Center Major Investigations Web site: www.nipc.gov.
5. National Infrastructure Protection Center Major Investigations Web site: www.nipc.gov/investigations/bloomberg.htm.
6. Sullivan, 2001.
7. National Infrastructure Protection Center briefing, August 2001.
8. http://www.info.gov.hk/police/aahome/english/statistics/download/200201/crimebrief_eng.doc
9. SANS Institute Alert, March 8, 2001.
10. Predictive Systems "Global E-review," August 2001. www.chron.com/cs/cda/story.hts/metropolitan/929311 .
11. First reported by www.securityfocus.com
12. A compromise is defined as access to a person's computer systems and databases without his or her explicit knowledge and consent. S1 had an impressive client list, from E*Trade to FleetBoston Financial Corp.
13. Sara Goo of the *Washington Post* first broke this story. www.idg.net.
14. National Infrastructure Protection Center. www.nipc.gov. These intrusions were perpetrated to steal proprietary databases, which were then sent to the heads of these banks with extortion demands.
15. The National Infrastructure Protection Center reported that the worm was distributed from unknown sources and is said to have disrupted and infiltrated networks worldwide. www.zdnet.com
16. U.S. District Court Arrest Warrant Case # 02-841.
17. U.S. Department of Justice, 2002.
18. www.newsbytes.com/news/02/175977.html.
19. The Economic Times http://economictimes.indiatimes.com/articleshow.asp?artid+14588

Just as legitimate use is increasing at a phenomenal rate, nefarious activity is also growing rapidly. Identity theft is the number one crime in the United States. Reported incidents of identity theft are projected to more than double, from 700,000[14] in 2001 to 1.7 million in 2005, and the costs to U.S. financial institutions alone will increase 30 percent each year, to more than $8 billion in 2005.[15] These numbers do not take into account the wide range of social costs associated with this crime, such as litigation expenses, or the lost hours to redeem one's name or credit information. In fact, these calculations do not include the very substantial losses for financial services providers generated by denial-of-service attacks. Table 3 suggests that denial of service can cost an average-size brokerage firm $6.5 million an hour or a credit card authorization company $2.6 million an hour. And these estimates do not include the costs of damage to reputation. Box 1 provides a graphic example of how pervasive a problem identity theft has become.

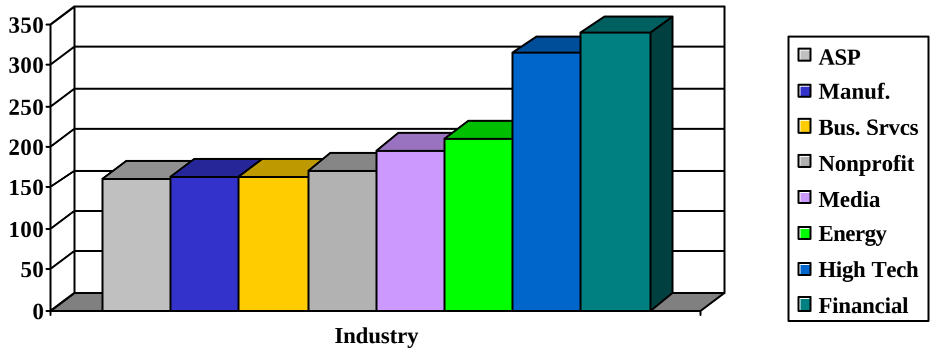
Table 3. Potential Losses from a Denial-of-Service Attack[16]

Business type	*Brokerage firm*	*Credit card authorization company*	*Automated teller machines*	*Major online auction site*
Exposure/Hour	$6.5 million/hr	$2.6 million/hr	$14,500/hr in fees	$70,000/hr

Source: Red Herring, December 2000.

Hacking, too, is endemic. Law enforcement agencies have documented that Eastern European organized hacker groups have penetrated hundreds of banks worldwide. The FBI's computer crimes division, the National Infrastructure Protection Center (NIPC), notes that many banks are paying off extortion demands for fear of risking their reputations and losing their customer bases to competitors. The Egghead hacking incident of 2001 represented a case of extortion. Hackers penetrated a database containing 10,000 credit card numbers and then demanded that the company pay them a large sum of cash to protect against the posting of those numbers in a chat room. Despite hackers' assurances to the contrary, every one of those compromised cards was charged a twelve dollar fee.

Figure 3. Hack Attacks in Asia (by Industry)



Source: Riptech Study 2001. ASP = application service provider.

[14] This figure represents a yearly trend within the United States only.

[15] Published in a 2001 report by Celent Communications. The projections were made using data from the Federal Trade Commission.

[16] Network shutdown.

As depicted in Riptech's recent study, the Financial Sector is the most commonly targeted industry in Asia. Recently, a fraud scheme that Australian police say may have netted millions of dollars, has shaken faith in one of the country's largest electronic banking systems. Almost 3 million Australians registered for Internet and telephone banking are being asked to check their accounts for irregularities following security breaches involving the Bpay electronic funds transfer system. Over 100 customers lost upwards of $150,000 each.[17]

Viruses are another computer-transmitted disease that swiftly compromises a system's integrity. A virus sets up residence in a system, and it is virtually impossible to kill it without replacing the infected parts of the system. Viruses did not exist before the early 1980s. Only recently have countries implemented legislation that makes infecting a system with a virus a crime.[18]

Box 1. Identity Thief: Abraham Abdallah

The most infamous of all identity thieves was Abraham Abdallah, a Brooklyn busboy. When police arrested him in March 2001, he had *Forbes* magazine's issue on the 400 richest people in America, as well as Social Security numbers, credit card numbers, bank account information, and mothers' maiden names of an A list of intended victims drawn from the issue, including Steven Spielberg, Oprah Winfrey, and Martha Stewart. Abdallah is accused of using Web sites, e-mail, and offline methods to try to steal the celebrities' identities as well as millions in assets. In May 2001, the Justice Department said in a statement to a congressional panel on Internet fraud, "Identity theft is the nation's fastest-growing white-collar crime." John Huse Jr., the Social Security Administration's inspector general, testified that the misuse of Social Security numbers in fraudulent activity is a "national crisis."

Source: New York Electronic Crimes Taskforce 2001

The banking system is no more vulnerable than the securities or the insurance industries. The U.S. Treasury recently discovered an infiltration of the electronic distribution system for its securities (see Box 2). In this case, defects in the risk-management processes employed by U.S. Treasury Direct in permitting access led to a situation in which one individual who was not creditworthy was almost able to compromise the whole system.

Both examples in Boxes 1 and 2 illustrate that the overall risk-management system permits a breach. Usually the breach does not result per se, from the adoption of a specific technology employed, but occurs because appropriate risk-management processes were not implemented.

Trends in cyber crime reveal significant growth. Attacks on servers doubled in 2001 from 2000. The 2002 CSI/FBI Computer Crime Survey[19] reported that 90 percent of organizations in the United States (including large companies, medical institutions, and government agencies) detected security breaches. Moreover, 70 percent in 2001 versus about 60 percent in 2000 reported serious security breaches such as theft of proprietary information, financial fraud, denial-of-service attacks, and compromising of networks. In most of these cases, the organizations cited their Internet connection as the critical point of attack. The 2002 CSI/FBI Computer Crime and Security Survey also indicated that 273 companies lost more than $266 million. Most important, according to U.S. law enforcement authorities, these numbers likely understate actual intrusions

[17] First reported in an article entitled "Australian bank scam may have siphoned millions." http://business-times.asia1.com.sg/news/story/0,2276,51947,00.html?

[18] Robert J. Morris wrote a computer program known as a worm that brought U.S. computers to an abrupt halt in 1988. Hackers break into California state computers. May 27, 2002, Associated Press.

[19] See http://www.gocsi.com/

and associated losses. When considered worldwide, these trends are even more troubling, given the relative sophistication of the U.S. security industry and the protections employed by financial services providers.

Box 2. U.S. Treasury Direct: The Case of Louis Lebaga

Louis Lebaga has been indicted for manipulating the Treasury Direct online system for auctioning U.S. government debt in order to obtain $1.3 billion of U.S. securities at the end of January 2002 in an attempt to defraud the U.S. Treasury and Union Bank of California (UBOC). He is accused of opening a deposit account at a bank and also establish a Treasury Direct account. He was able to make what amounted to a very large number of bids for U.S. Treasury bills amounting to $160 million and 231 bids for U.S. Treasury notes totaling $1.155 billion between January 29 and January 30. He could do this because the Treasury Direct system provides for no explicit cutoff with respect to the number of bids that one retail investor can make, and it does not require the explicit posting of cash and collateral. On the first issue date for the securities on which he bid, Lebaga was able to get $160 million in securities delivered electronically by the Federal Reserve (the fiscal agent for the Treasury) to a bank account at the UBOC and to have funds transferred electronically to pay for the purchases. After UBOC realized that good funds were not in the account, it notified the Federal Reserve, and the funds were returned the following day. Note, however, that if Lebaga had not wanted to carry out more transactions, he would have had more time and options to transfer funds or services.

Further, Lebaga was able (because of human error and defects in the U.S. Treasury Direct system for checking documentation of users) to obtain this electronic transfer of funds without adequate credit checks. On arresting Lebaga, law enforcement authorities learned that he was $2,000 in arrears on his rent and that eviction papers had been served. The reality is that this one individual could have used these techniques to effectively compromise the Treasury Direct system. It is interesting to note that both Brazil and the Philippines have developed electronic distribution systems for government securities, catering to retail investors. The adoption of insecure risk management frameworks will eventually their integrity as well.

Source: John Frazzini, Special Agent for U.S. Secret Service Financial Crimes Division. Interview on April 8, 2002.

In the United States, a 2001 CSI/FBI Computer Crime Survey identified the following five major reasons organizations did not report electronic intrusions to law enforcement:

- Negative publicity.
- Negative information competitors would use to their advantage——for example, to steal customers.
- Lack of awareness that they could report events.
- Decision that a civil remedy seemed best.
- Fear among IT personnel of reporting incident because of job security.

Lack of accurate intrusion reporting to regulators and law enforcement is the core reason that issues related to electronic security are not being recognized as an immediate priority.

III. The Electronic Security Industry

Today's electronic security industry boasts an ever-growing array of companies. The types and numbers of choices can be confusing for the expert and overwhelming to the novice. These companies are involved in every facet of securing the networks used by financial services providers. They range from those that provide active content filtering and monitoring services

[20] See http://www.gocsi.com/

(even virus detection companies are an example) to those that undertake intrusion detection tests, create firewalls, undertake penetration testing, develop encryption software and services, and offer authentication services.

In scope, the e-security industry increasingly is becoming a worldwide presence as it grows parallel with the expanding connectivity to the Internet. The growing integration of technologies among the Internet, wireless, Internet provider (IP), telephone, and satellite will also present new challenges for electronic security and the structure of the financial services industry and e-finance.

From the vantage point of financial services providers, the earlier that security is built into a system's design process, the greater will be their return on investment in security-related services. For example, studies show that spending $1 to fix a vulnerability during the system design process saves $99 of the $100 that must be spent later when the system is implemented (See Berinato 2002; Soo Hoo 2001). This cost avoidance or cost savings makes or breaks many IT projects. The increasing extent to which technology platforms drive financial services and the increasing rates at which computer electronic security incidents are occurring emphasize the importance of using risk management in making business decisions to avoid greater future losses.

Electronic Security Vendors

A rich variety of vendors operate in what is becoming a global industry for electronic security. Many types of companies operate in this industry. In the United States alone, $5.1 billion in security software was sold in the year 2000—a 33 percent increase over the prior year.[21] These companies are involved in every facet of securing the wide area networks over which financial services are provided. The following is a brief description of the major categories of vendors. (See also Figure 3.)

Active Content Monitoring and Filtering.[22] Companies involved with active content monitoring and filtering produce tools that examine for potentially destructive content material entering a network. These vendors provide tools to monitor all content entering a network for malicious codes, such as harmful attributes. Trojans, worms, and viruses are methods used to deploy an attack once the perpetrator enters the system. Viruses are programs that infect other programs on the same system by replicating themselves. Virus scanners are critical in mitigating these attacks. Vendors of virus scanners provide software that scans and cleans networks and is periodically updated.

Intrusion Detection Systems Vendors.[23] Companies that produce network intrusion detection systems provide products to monitor network traffic and alert the systems administrator with an alarm when someone is attempting to gain unauthorized access.

Firewall Vendors.[24] Companies that produce firewalls provide a virtual "security guard" at the gate of the customer's facilities. A firewall is a system that enforces the access-control policy between two networks. Vendors create these virtual guards to protect a network's integrity.

[21] See Cunningham, "Digital Security: Heightened Risks Demand Innovation," Red Herring, July 2001.
[22] For more details on this facet of the industry, see Annex II.
[23] Ibid.
[24] Ibid.

Penetration Testing Companies.[25] These consulting organizations simulate attacks on networks to test for a system's inherent weaknesses. They then patch the holes found during the simulated attacks. Typically, vulnerability-based scanning tools provide a current snapshot of a system's vulnerabilities.

Cryptographic Communications Vendors.[26] Vendors who supply this product enable the client company to protect its communications with an encryption envelope. Encryption uses complex algorithms to shield messages transmitted over public channels. It provides safe passage from point A to point B. When the message reaches its destination, the recipient uses another algorithmic key to open it. It is highly recommended for use by mobile workforces and/or large noncentralized corporations or institutions.

Authentication Vendors. Authentication asks users such questions as "Who are you?" and "Are you allowed to do that?" and permits a user to access the system only if these questions are answered correctly. This type of service can be broken into four general categories: passwords, tokens or smart cards, biometrics, and encryption. (See Annex I for more details.)

Links to E-Finance

Because E-security companies are becoming increasingly global in nature, it is important when designing public policy to understand the links between such companies and the electronic finance industry. Figure 4 provides a stylistic example of some of the links among the many types of vendors of electronic security services and financial services providers.

Figure 4 also shows a potentially disturbing reality about the electronic security industry. One vendor may provide multiple services to several interlinked customers. For instance, a vendor may provide security to the financial services provider's online platform. This same vendor also may provide security services directly to the bank for its offline computer systems. In addition, it may supply security services to the hosting company. Telecommunication companies in many emerging markets provide hosting—or what many refer to as "e-enabling services"—to the banking community. By establishing a convenient online platform that customers can access through a variety of electronic devices, these hosting companies (e.g., ISPs) have become targets of organized crime.

In many emerging markets, the telecom company may have an interest in or own outright the ISP provider and the hosting company and may provide various forms of financial services as well. Moreover, many telecom companies also have multiple interests in many different forms of technology providers, from fixed-line telephony to wireless to satellites. This industry structure should raise concern—it signifies the need to discuss and debate difficult public policy issues now, such as competition policy, and how these issues might be addressed in designing new legal and regulatory elements of the present frameworks (see Claessens, Glaessner, and Klingebiel 2002).

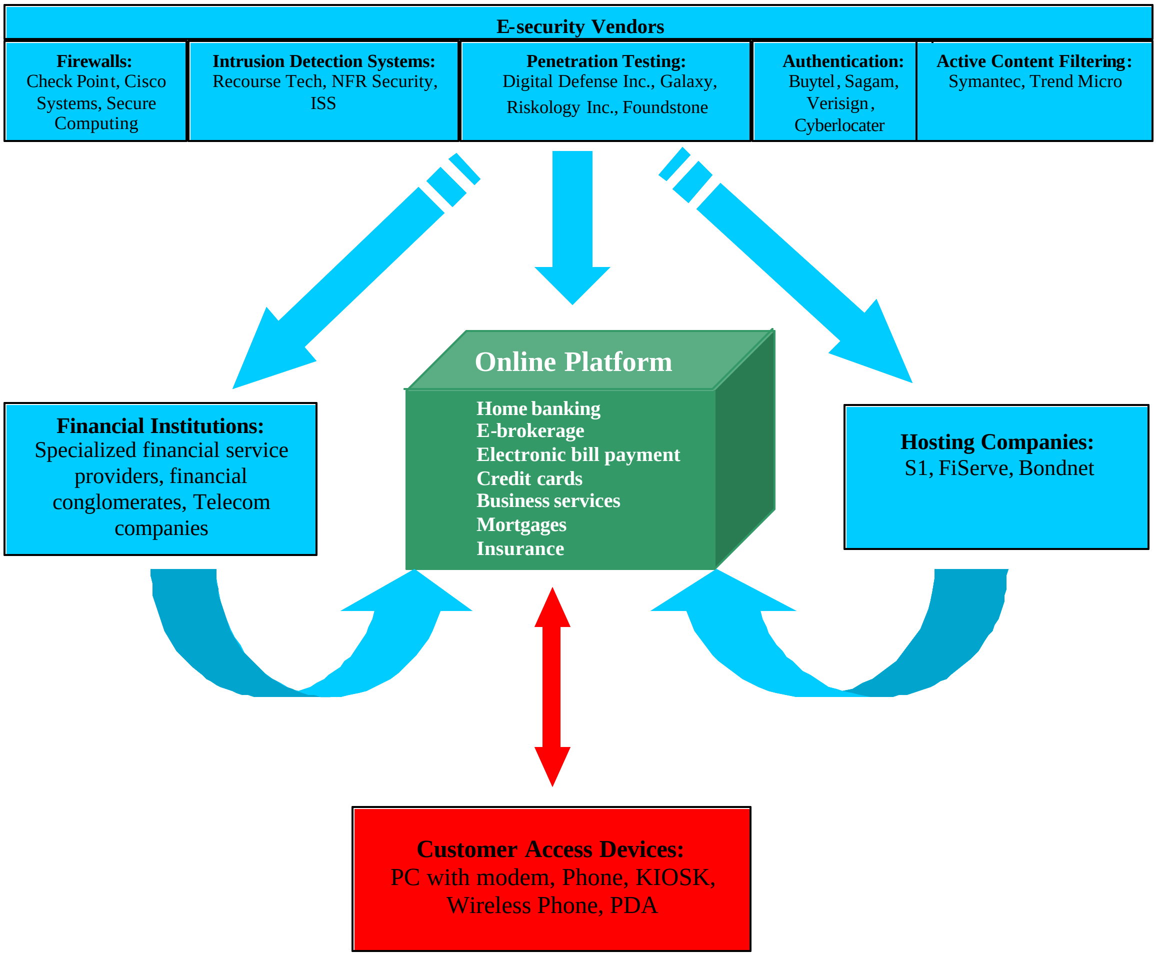
Along with a complex industrial organization, convergence in technologies will present special challenges in the design of public policies relating to electronic security. Specifically, increasing points of vulnerability will exist, and any well-designed electronic security system must address them. These new points of vulnerability might include the potential interfaces between customer access devices, such as a PC with modems, land-line phones that can be linked

[25] Ibid.

[26] Ibid.

with any Internet platform through voice recognition, wireless phones, or personal digital assistants (PDAs) with an online platform. The point at which the message leaps from one channel to another is the point at which it is most vulnerable. Hence, financial services providers will need to address a much wider array of risks and expend effort to define liability, and public policymakers will need to examine the impacts of potential weaknesses, given what is already a complex e-finance industrial structure.

Figure 4. E-Security Industry and E-Finance


E-security Vendors
Firewalls: Check Point, Cisco Systems, Secure Computing
Intrusion Detection Systems: Recourse Tech, NFR Security, ISS
Penetration Testing: Digital Defense Inc., Galaxy, Riskology Inc., Foundstone
Authentication: Buytel, Sagam, Verisign, Cyberlocater
Active Content Filtering: Symantec, Trend Micro
Online Platform
Home banking
E-brokerage
Electronic bill payment
Credit cards
Business services
Mortgages
Insurance
Financial Institutions: Specialized financial service providers, financial conglomerates, Telecom companies
Hosting Companies: S1, FiServe, Bondnet
Customer Access Devices: PC with modem, Phone, KIOSK, Wireless Phone, PDA

Box 3 highlights an inherent conflict: The need to secure systems against physical risks that can involve use of multiple technologies in different locations runs up against the fact that the most distributed and decentralized networks are more vulnerable to interception and unauthorized access at the point of interface. As technologies converge, development of more effective standards for securing such points of interface will become far more important.

One example of how convergence of technologies creates vulnerability occurs when a wireless Groupe Spécial Mobile (GSM) phone is used to initiate a transaction through an interface with the Internet (e.g., via indicating transactions on the online platform of the financial services provider). Specifically, a secure way of integrating between the two technologies—GSM and the Internet—is needed. This typically requires seamless connectivity and an integration of standards, including those for security worldwide, that are not in place today. Wireless messages have to travel through a gateway,[27] which channels them to a wired network (e.g., the Internet) for retransmission to their ultimate destination. At the gateway, the message sent and encrypted in GSM using what is called Wireless Application Protocol (WAP) and the associated use of Wireless Transport Layer Security (WTLS) must be converted into the industry standard for secure messaging over a wired network—secure socket layer (SSL). At this point (in the gateway), the message will be unencrypted before being reencrypted, and there is vulnerability.

Box 3: Evolution of Technology and International Standards

A fundamental security dilemma is embedded in global information networks. The global IT network infrastructure (i.e., the physical machinery that allows electronic linkages on a global scale) is on the one hand most secure and immune to destruction by natural disasters or terrorist and cyber-criminal attacks when it is most distributed and decentralized. After the great Kobe earthquake and the destruction of all fiber optic connections in the city during that natural disaster, the Japanese government proceeded to install satellite VSAT[28] terminals at post offices throughout the country to ensure the basic integrity of national communications.

On the other hand, the most distributed networks are most vulnerable to interception and unauthorized access. Often maximum vulnerability to interception exists at the point where fiber, coax, satellite, and terrestrial wireless systems interconnect. Air interface standards are but one example of modern telecommunications and IT systems open to interception.

A few years ago, telecommunications were projected to follow the model known as the "Negroponte Flip," whereby all narrowband traffic would go onto wireless (and largely mobile) systems, but all broadband service (in order to conserve frequency) would go onto fiber networks. This model was focused on the United States rather than the rest of the world, and it was technology-driven. In fact, the popularity of wireless systems (both satellite and terrestrial wireless) has continued to increase, and the market has demanded more and more Direct Broadcast Satellite (DBS) satellite entertainment service and broader band wireless now in the form of third generation systems and soon fourth generation systems. The market has thus actually followed the trend of the so-called Pelton Merge, which calls for continued improvement of "seamless interface standards" that allow the smooth interconnection of fiber, coax, terrestrial wireless, satellites, and other new and evolving technologies such as high altitude platforms. The challenge is thus to develop standards that allow easy and reliable interconnection and yet also protect security.

One example of a standard is the ISO seven-layer model of telecommunications. The current standard, however, does not really treat the security issues in the seamless interface between these technologies. Hence, it is necessary to consider the creation of a new layer that is truly secure, based on a 256- or even 1024-bit code that is constantly updated. Further study would be needed to determine whether the ultimate solution is a separate layer or the reengineering of part of an existing layer that could be devoted to this task. The extension of security identification module (SIM) smart cards that could be used throughout the world would also be a major step forward.

Source: Dr. Joseph Pelton, Executive Director of the Clarke Institute

[27] For more detailed analysis of this problem, see Annex III.

[28] **V**ery **S**mall **A**perture **T**erminal, but more simply put it describes a small satellite terminal that can be used for one-way and/or interactive communications via satellite.

IV. Electronic Security Infrastructure in a Risk-Management Framework

Regulation of the Electronic Security Industry

To develop a framework for thinking about the public policy issues that arise in examining electronic security, it is necessary to identify the fundamental source of "public interest" and the case for regulation in this area. For several reasons, electronic security warrants some form of public intervention now.

First, financial services and the payment system in particular, or banking more broadly, constitute one of the eight identified areas of "critical infrastructure."[29] A compromise of the payments system caused by illegal access and hacking can have broad ramifications for an entire economy, as could similar impacts in other critical infrastructure areas, from transportation to energy, and so on. Hence, the public interest and welfare are potentially at risk when business and commerce fail to meet certain minimum electronic security standards.

Second, the role of government and law enforcement can be justified on much more familiar classic market-failure grounds.[30] Specifically, the existing base of information that supports projections about the extent of the electronic security problem is substantially flawed. This is because financial services providers, hosting companies, and other enabling companies have inadequate incentives to report intrusion or penetration information accurately, given their legitimate concerns about the disclosure of such information and its potential damage to both their reputation and public confidence in their business. In this case, insurance markets cannot price the insurance risk in an actuarially fair manner. Similarly, information technology is subject to large increasing returns to scale on both the demand side and the supply side (see, e.g., Shapiro and Varian 1999). Market outcomes in such industries (including financial services, which is heavily dependent on IT) will tend to be somewhat concentrated and often will require industry standardization and coordination.

Any approach to defining public policies through law and regulations (including prudential regulations, such as capital standards) must account for the impacts electronic security considerations or the lack thereof have on a set of risks. Specifically, financial services providers react to incentives. In many cases, analysts pressure financial services providers to produce targeted returns, while at the same time pushing them to outsource in order to reduce costs. Meanwhile, technological advances have created a much more complex inter-relationship between electronic security and risks of different types. In effect, electronic security and electronic finance can have an impact on operational risk, risk of identity theft, fraud and extortion, credit quality deterioration, and systemic risk, and can even have implications for the risks in undertaking failure resolution.

[29] The Policy on Critical Infrastructure Protection: Presidential Decision Directive 63 (PDD-63), issued by the Clinton Administration in 1998, provided a starting point for addressing cyber risks against the United States. This directive identified the critical sectors of an electronically dependent economy and assigned lead agencies to coordinate sector cyber-security efforts. This directive identified eight sectors—finance, transportation, energy, water, government, aviation, telecommunications, and emergency—presenting the vision that "the United States will take all necessary measures to eliminate swiftly any significant vulnerability to both physical and cyber attacks on our critical infrastructures, including especially our cyber systems."

[30] Classic reasons for a failure in a market are asymmetric information, increasing returns to scale, and network externalities. See Bator (1967), Varian et al. (1999), and Kahn (1970).

Operational Risk. Inadequate electronic security can result in interruptions of service and—in some cases, depending on the nature and adequacy of backup systems—even the loss of critical information. As part of managing operational risk, financial services providers worldwide need to pay greater attention to the way they secure their IT systems. As discussed in Section VII, the risks involved in electronic security often relate to extortion and reputation risk, which usually are not specifically taken into account in the allocations set aside to cover operational risk.

Risk of Identity Theft, Fraud, and Extortion. As noted in Section II, penetration by hackers often leads to extortion demands. In addition, identity theft is a growing concern for e-finance service providers. Its growth has been rapid, but as in the case of hacking, it is not reported in a timely manner or accurately; thus, its growth may be considerably understated. This problem is not unique to financial services—it also affects the integrity and reliability of the credit information gathered and assessed by credit bureaus, downstream to credit decisions.

Risk of Credit Quality Deterioration for the Financial Services Provider. Although not often acknowledged, a substantial denial of service or long-term intrusion that results in fraud, impersonation, or corruption of data can effectively cripple a bank's operations for a period of time. If that time is sufficient, it can irreparably damage the bank's reputation and possibly compromise its credit standing. Because market participants' confidence is critical, such an event could have a pernicious impact in a relatively short time.

Systemic Risk. One of the most important links between e-finance, e-security, and risk is the systemic impact that the associated risks can have on the related payment systems through interaction with compromised networks. Appropriate security should be proportional to the value of underlying transactions. For this reason, in the case of large-value clearinghouses, extensive electronic security is or should be in place. Any intrusion or interruption in a payment system's electronic messaging could easily create significant system-wide exposure.

Risks in Failure Resolution. A final form of risk associated with the delivery of e-financial services and security relates to the risks introduced when a brick-and-clicks or wholly Internet-based bank fails. Here the process of closure itself is difficult to define and even more difficult to implement if the entity has its servers in offshore centers. Closure in this case would require extensive cross-border coordination among authorities in what could be numerous disparate jurisdictions. Cooperation, and thus closure, may not be feasible with the speed that can be applied in the case of a non-Internet-based bank. At the point of intervention, if the records and other essential information about digital assets are not preserved under well-defined guidelines, and if they are not secured or cannot be retrieved from servers, then, at the very least, claimants' rights may be compromised.

Trade-Offs: Security, Quality of Service, Privacy, Technological Innovation, and Costs

Designing public policy in this highly complex area requires balancing a number of essential trade-offs in creating legislation and regulation. This even applies in designing standards and guidelines that might be used by a self-regulatory agency or by an official agency.

Security and Costs. Security should always be proportional to the real value of the underlying transaction. Given this proviso, it appears that when transaction value is small, no clear economic or risk-management case can be made for employing the most sophisticated electronic security regimes when a less expensive form of security will yield the same return. For example, a financial services provider would not want to use an expensive and cumbersome authentication process, such as public key infrastructure (PKI), for small-value transactions when tokens or other simpler forms of authentication will mitigate the risk of theft, and so on, to an acceptable level.

Security and Quality of Service. Similarly, trade-offs exist between the convenience or quality of service, as computed in terms of speed and the extent and degree to which security is used. The more complex the security process used, such as PKI, the longer the transaction takes to be completed. Advances in these technologies are lessening this trade-off. Over time, effective authentication or encryption systems will be available that do not slow the speed of transactions and do not disparage the quality of service. Moreover, one can argue that confidence in the security of services is an essential aspect of quality in providing financial services.

Security and Technological Innovation. For electronic security systems to be effective, it is important to ensure that private parties agree to certain standards and guidelines. But the proliferation of technologies that can be used to transmit information and their rapid rate of integration inherently creates a reluctance to adopt standards or guidelines. Technological innovation can be stifled and customer service can suffer if security standards are not sufficiently flexible and technology-neutral. As will be noted in later sections, even the definition of an electronic signature needs to be very carefully designed so as not to preempt the use of a number of alternative technologies. In other words, the concept of technology neutrality is an important one to adopt when formulating legislation and regulation. (See Section VI.)

Security and Privacy. Ironically**,** the need for more effective electronic security may sometimes conflict with and negatively affect the user's privacy. Inadvertently, it may also affect the privacy of third parties who are identified in affected information. This tension is natural, and it is not new. On the one hand, certain types of electronic security services may be consistent with protecting privacy (e.g., programs such as cyber patrol). On the other hand, security may be needed to track and verify the user's movements. In other cases, however, the person undertaking the transaction may want to remain anonymous as part of a trading strategy. Developing the proper balance between security and privacy is a delicate matter. It often is decided within a cultural paradigm. Sometimes this means that something considered private in one culture may not be deemed so in another. Moreover, the laws (e.g., bank secrecy provisions) often compromise the ability of the authorities to investigate properly and take enforcement actions in complex electronic crime cases.

The Pillars of an Overall Framework

This paper is built on the concept that trust and confidence of market participants is a key component of a robust economy. Given this assumption**,** seven fundamental pillars are needed to sustain a framework of reform and to improve the security of the market. These are

- An adequate legal and enforcement framework in certain critical areas.
- Adequate treatment of electronic security in the case of payments services and those that undertake to provide e-enabling services to financial services providers, such as money transmitters.
- An effective supervision and prevention regime to manage emerging electronic security requirements.
- Public partnerships with private insurance companies to monitor the efficacy of security systems on a macro level and promote the development of minimum standards for electronic security.

- Public partnerships with private entities to develop and adopt transactional security levels for transactional information and electronic signatures, together with criteria to protect document and data classification standards.
- Public partnerships with private entities to develop and maintain accurate incident databases and a related reporting framework for electronic security incidents to be used in assessing systemic risk over time.
- Public education about how technological change and related electronic security risks need to be addressed.

Issues usually arise in each of the areas identified above when the challenges posed by electronic security are addressed in a more systemic manner. The sections that follow explore each of these pillars.

V. Pillar I: Legal Framework and Enforcement[31]

Laws, Policies, and Practices Bearing on Electronic Security

Countries adopting electronic financial services should address and incorporate security concerns as they develop policies, laws, and regulations. In this way, they can build a security framework that will support the safe and sound operation of their institutions and combat crime and cyber terrorism. The following areas of law, at a minimum, should be included in any e-finance legal framework:

- Electronic transactions and commerce law
- Payment systems security law
- Privacy law
- Cyber crime law
- Anti–money laundering law

These five categories of law address the *basic relationships and transactional activity* that flow through the e-payments system.

The cornerstone of an e-finance legal framework is recognition of the legal validity of electronic signatures, transactions, or records. Further, these laws should prefer technology-neutral solutions, provide basic consumer protections for electronically based transactions, promote interoperability, and address records retention. Two basic models exist: the act developed by the United Nations Commission on International Trade Law, titled UNCITRAL, and the Uniform Electronic Transactions Act (UETA). An electronic commerce law might address all non-consumer-related financial transactions and records. It should focus on governing conduct with consumers and on basic financial payment mechanisms such as EDI, EBT, EFT, and ETC. Specifically, it defines what constitutes a secure financial services system in an open network architecture and requires entities to practice due diligence.

Electronic Transactions and Commerce Law: The past seven years have produced tremendous growth in electronic-commerce-related legislation. In 1995, only a handful of

[31] The authors thank Edward Gilbride, Counsel for the Federal Deposit Insurance Corporation, for very helpful written inputs in the context of the discussion in this section. Thanks are also owed to John Carlson, Senior Adviser/Deputy Director, Bank Technology Division, OCC.

countries had basic computer or intellectual property laws. Today, almost every country has enacted an electronic signature or electronic transaction act. The basic elements of these laws are the same, with minor variations. Most of the laws use UETA, promulgated in the United States by the National Conference of Commissioners on Uniform State Laws (NCCUSL), or UNCITRAL.

Significant differences exist in the provisions of UETA and UNCITRAL, but the objectives of both are the same: to promote electronic commerce and to ensure that electronic signatures, however they may be defined, have the same effect under the law as manual signatures. For example, UETA defines an electronic signature as "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." UNCITRAL[32] defines an electronic signature as "data in electronic form in, affixed to, or logically associated with a data message, which may be used to identify the signatory in relation to the data message and indicate the signatory's approval of the information contained in the data message." Each provides a different perspective on timing and intent. UETA presumes that by signing the document, the signer intends to be legally bound. Its wording creates a presumption in favor of the validity of the contract. UNCITRAL, in contrast, uses permissive language, creating no presumption in favor of the contract. Further it should address the issues of record management and record retention.

With the proliferation of electronic signature and electronic transaction legislation over the past decade, electronic commerce has come into its own legally. In general, an electronic signature has the same force and effect as a manual signature in most of the world. The latest country to adopt electronic signatures was Russia, which enacted its Electronic Digital Signature Law on January 16, 2002. Typically, the law changes significantly more slowly than many other parts of a culture. The law appears, though, to be trying to adapt to electronic commerce needs as quickly as the world is coming online. This is a major phenomenon that raises issues of importance beyond the scope of this paper.

Payment Systems Security Laws. Though most countries have laws in place to regulate different components of the payments system, no country has yet addressed payments systems issues comprehensively. Payment systems legislation should identify, license, and regulate any directly related payment system entities, such as money transmitters and ISPs. It should require such elements to operate in a safe and sound manner so as to protect the integrity and reliability of the system. It should require the timely and accurate reporting of all security incidents, including all electronically related money losses. Finally, it should require all payment system entities to adhere to a documented security program and should encourage some form of shared risk protection.

Privacy Laws. Clearly, privacy is an area of the law that is undergoing considerable scrutiny throughout the world. It is an issue of fundamental importance, reflecting the very substance of our cultural identities, values, and mores, and it must be handled with the utmost care. Poorly considered decisions made in this arena may haunt us for years to come.

On the issue of privacy protection, some countries have chosen to legislate on a functional or piecemeal basis, while others have taken a more encompassing, process-oriented approach. Two approaches are also being used on the issue of consent. The first is to assume consent unless the party affirmatively chooses not to have the information sold or used for other purposes. The second is to assume that the party has not consented to any use of the information unless the party gives that consent. The United States follows the first approach in financial

[32] Article (II) Subsection A of the 2001 UNCITRAL Model Law on Electronic Signatures.

activity and the second in medical information. The European Union (EU) exemplifies the second in each area and continues to be the leader in providing privacy protection to its citizens with its 1990 EU Directive on Data Collection.

No matter which approach is used, at a minimum, privacy laws should embrace the Fair Information Practice Principles of notice, choice, access, and security. They should address privacy rights concerning any data collected, stored, or used by an entity for different purposes, in particular those uses that could affect a person's basic human rights, such as criminal, financial, business, or medical uses. In practice, privacy laws would require entities to do the following: advise persons about how data will be used; collect only the minimum data needed to complete the transaction or record at issue; use the data only for those purposes that it advised the person it would be used for; and permit persons to view any information collected and dispute the validity of any such information with timely corrections. Finally, the law should impose restrictions on any entity collecting, holding, or disclosing information in a form that would allow identification of the person it relates to, however that may be defined.

In practice privacy laws would require information gathering entities to advise persons from where they are collecting the information and how the data will be used.[33]

Cyber-Crime Laws. Significant debate is transpiring in legal communities worldwide over the impact of cyber crime on fundamental concepts of law, such as jurisdiction, and in particular on how the electronic culture is changing traditional paradigms. Financial cyber crime is a top priority in this dialogue because, more often than not, it requires intense international cooperation among what can be an overwhelming number of law enforcement agencies and regulators from different countries. Because no country is immune, every country should benefit from pooling resources to address this problem. But, more than any other aspect of computer law, financial cyber crime tests the continuing validity of the industrial regulatory and law enforcement model. As a result of their lack of Cyber-crime legislation the Ukraine and Belarus have become major staging grounds for organized hacker syndicates. Because of the underlying complexity of such cases and the overlapping jurisdictions of authority within a country, one of the first things the laws should address is who or what has authority and responsibility for these cases. A significant cost avoidance could result from such reform, and money saved could be invested in trained resource experts and the tools needed to investigate, prosecute, and punish cyber-crime perpetrators. Substantively, the laws should address abuses of a computer or network that result in loss or destruction to the computer, the network, or people, and should include provisions for restitution for associated losses.[34]

A December 2000 McConnell International survey provides a snapshot of the state of computer crime legislation worldwide. It examined the legal frameworks of 52 countries to determine each one's ability to prosecute perpetrators of 10 types of computer crime. The survey showed that a patchwork of outdated and inconsistent laws effectively functions as a shield from prosecution for cyber criminals who attack electronic systems and information.[35]

[33] This data should be used only for those purposed that were intended. They should also permit the persons from whom they collected the information to view it and provide a process by which, such persons could dispute the validity.

[34] The United States has enacted various computer intrusion laws that treat identity theft and computer-initiated fraud as criminal offenses with severe penalties. Recent legislation grants individual banks the power to freeze customer accounts if criminal activity is suspected. Penalties for fraud and related activities perpetrated in connection with computers can include imprisonment of up to 25 years (see http://www.cybercrime.gov/cclaws.html).

[35] See http://www.mcconnellinternational.com/services/securitylawproject.cfm

For countries looking to develop cyber-crime legislation, the Council of Europe provides some guidance. In 2001, it developed the first international treaty on crimes committed via the Internet and other computer networks, dealing particularly with infringements of copyright, computer-related fraud, child pornography, and violations of network security. The treaty also provides for a series of powers and procedures, such as the search of computer networks and interception.[36]

Anti–Money Laundering Laws. Worldwide, money-laundering is recognized as one of the most potent forces threatening political and economic stability. Since 1990, the Financial Action Task Force (FATF) has spearheaded the adoption and implementation of measures designed to counter the use of the financial system by criminals (see http://www1.oecd.org/fatf/). It established 40 recommendations that set out the basic framework for anti–money laundering efforts and are intended to be of universal application. In 1996, the FATF recognized the link between cyber vulnerabilities and money laundering when it modified its 40 recommendations 1996 to include number 13, which states, "Countries should pay special attention to money laundering threats inherent in new or developing technologies that might favor anonymity, and take measures, if needed, to prevent their use in money laundering schemes." The points addressed in cyber-crime laws also apply here. Substantively, at a minimum, these laws should define money laundering and should commit to international cooperation in the investigation, prosecution, and punishment of such crimes pursuant to the guidance provided by the FATF. The FATF regularly reviews its members for compliance with the 40 recommendations, with the result that the recommendations are now the principal standard in this field.

Enforcement Powers

The ability to enforce the laws and regulations within and across boundaries is as important as providing an adequate legal and regulatory framework within which to prosecute perpetrators and penalize those entities operating in an unsafe and unsound manner. To achieve enforcement, many countries need to take a number of critical steps.

Regulatory enforcement reforms should address, at a minimum, varying degrees of cease-and-desist orders and compliance actions. Cease-and-desist orders could range from removal of the entity from the online system until it comes into compliance to closing the entity down. While a financial services provider may not have access to online activity, it still may be conducting unsafe and unsound operations to such an extent that it is jeopardizing other entities.

Without a concerted international cooperative effort, e-finance hackers will commonly move to jurisdictions with the most lax legal and enforcement frameworks.

Access, availability, and interoperability should be the mantra to guide financial supervision and enforcement efforts. The traditional regulatory structure must expand to include all entities directly related to the delivery of financial services. This entails everything from ISPs to ASPs, software and hardware vendors, and security providers.

Legislation needs to incorporate these providers into the regulatory and enforcement net. Moreover, professional liability needs to attach to these providers, to the directors who contract with them, and to the lawyers and accountants who provide services to them because, in the new paradigm, all are indispensable to the institution's ability to provide financial services. One

[36] See http://conventions.coe.int/Treaty/en/Treaties/Html/185.htm).

approach might require that these providers be bonded, licensed, and subject to periodic audits and examination under the appropriate regulatory scheme. This would create a relevant basis from which to undertake enforcement actions. As stated already, traditional regulatory schemes are outdated, and as currently configured they cannot adequately address the new components of the payments system to determine whether a financial institution is operating in a secure manner.

VI. Pillar II. Electronic Security of Payment Systems

Money Transmitters: Background

Convergence and integration are the keys to the revolution in money movement and to wholesale and retail payments services. Convergence of the telecommunications, computer, and financial services industries is changing the fundamentals of the industrial organization of the financial services sector, as noted in Section III, redefining traditional boundaries and jurisdictional limits of responsibility because of shifting legal, regulatory, and financial concepts. The industrial age gave rise to certain agreed-upon regulatory concepts by which the telecommunications and financial services industries operated. Regulation of telecommunications was based on public safety, interest, and welfare through the use of universal access and service. The regulation of banking was based on safety and soundness with nondiscriminatory access to credit opportunities. Convergence, however, requires reassessment of this regulatory paradigm because of the necessity for *universal access in a safe and sound environment*.

Convergence and integration help realize the telecommunication and financial services goals of access, availability, and interoperability. Access to the financial system was once limited to a few complex protocols. Now anyone can access the system using microwave, wireless, satellite, public switched network (PSN), computer, IP telephony, interactive television, ATM, or brick-and-mortar structures. In addition, these advances have redefined and eliminated time so that the financial system is accessible to anyone, anytime, anywhere, using cash, debit card, check, credit card, stored value card, or smart card. Money is now interoperable, as telecommunications and computers facilitate the conversion from one currency to another simply by the push of a button. Eventually, even the servers of a telecommunications company, in addition to facilitating cellular calls, will be used for effecting payments between prepaid cell phone subscribers.

Under this new industrial structure, and given the increased outsourcing of operations, the following questions about the design of regulations seem reasonable: Who or what is a money transmitter? What is an ISP? Should the regulatory framework deal only with core financial activities, or should it include outsourcing entities? If outsourcing increases, what is the case for regulating these independent entities, and what agencies need to play a role or have ultimate responsibility? Such fundamental questions must be answered to create effective incentives for money transmitters and ISPs to adopt adequate electronic security. The regulatory objective must be clear and simple.

Who or what is a money transmitter? Today, the set of entities involved in money transmission or payments is more difficult to define than one might expect. These entities are not well regulated or supervised in many countries, even if they can be defined. For the purposes of this paper, a money transmitter is any commercial enterprise that engages in the transfer and exchange of monetary instruments and currency.

Money transmitters may perform a variety of services, including money order issuance, wire transfers, currency exchanges, check-cashing, and check-presentment. More recently, money transmitters have been providing electronic check-presentment services and point-of-sale money payment order information to the accepting bank. Money transmitters operate outside the depository institution but are often are associated in some way with one or more depository institutions in a downstream relationship.

What is an ISP? Whether an entity is an ISP can be difficult to determine under existing law. ISPs are not regulated in most countries, and countries that have tried to regulate them have experienced significant backlash. One recent example involved Australia's Broadcasting Services Amendment (Online Services) Act 1999, referred to by its critics, who claim it is overly broad, as the Internet Censorship Act. It has received international attention and is touted as an attempt by one country to impose a censorship regime on the Internet.[37] A number of entities, including financial services providers, could fall under its definition of an ISP.

This paper suggests that analysis of the payment system at large shows that money transmitters hosting companies/ISPs have become a critical sector and can have a direct impact on the security of a financial service provider. As an example, the use of multiple channels to distribute financial services or make payments has expanded the circle of providers to include a Web site hosting service, a third-party software developer to plan and implement the Web site, application software or service providers, a third-party processor to facilitate information movement from the Web site to the financial institution's network, a customer service call center, and one or more ISPs or money transmitters. Use of these new channels means that the financial services sector now broadcasts; publishes; provides or uses e-mail, Internet services, network services, and entertainment; hosts online forums; and uses bulletin boards. As the nondepository institutions involved become more varied, defining who is a money transmitter becomes more complex and requires a two-part test. First, to what extent is an institution relying on that provider to transact and deliver financial services? Second, to what extent can the provider have an impact on the payment system?

The expansion of the types of entities involved in money transmission creates both greater opportunities and more complex liabilities and responsibilities. Converging technologies have opened access to the payment systems. Disintermediation of the financial services sector has created an open competitive environment to all aspects of the payments system. Open access has resulted in the proliferation of money transmitters and their partnering with ISPs. With these developments, challenges have increased for electronic security of payments.

Safety and Soundness for Money Transmitters and ISPs

The question of how to ensure safety and soundness in the case of ISPs and money transmitters must address at least five basic, generic problems:

1. Lack of definition,
2. Lack of reporting requirements,
3. Limited or no regulation,
4. Limited or no warranties, indemnification, and liabilities, and

[37] The Online Services Act defines an ISP as anyone who provides an Internet carriage service that is used for (a) the carriage of information between two end-users outside the "immediate circle" of the supplier, as defined in the Commonwealth Telecommunications Act of 1997—and when one person uses an Internet carriage service to view the content of a second person (e.g., by visiting a Web site), both of these people would be considered end-users of that carriage service; or (b) the carriage of information simultaneously to more than one end-user, at least one of whom is outside the immediate circle of the supplier.

5. Lack of security as a necessary element for service providers.

Toward a Working Definition of a Money Transmitter

Money transmitters are often referred to as nonbank financial institutions or money services businesses. Numerous definitions exist for this payments system "service" sector. Generally speaking, and for the purposes of this paper, money transmitters are commercial enterprises engaged in the transfer and exchange of monetary instruments and currency. In the context of electronic payment systems, they typically serve as third-party automated clearinghouse (ACH) providers.

Money transmitters do not operate alone. They require access to telecommunications to transport information from point to point. Usually a money transmitter contracts with an ISP to transport the information across network lines.

Failure to require reporting or to review and expand regulations to include new money movement vehicles permits unsafe and unsound activities to use the payment system without check or prevention. Legislation should place an affirmative duty on executives to report incidents, and the intentional failure to report should carry potential punishment.

Liability of Money Transmitters

The money transmitter-ISP venture is usually structured as a layered relationship built on successive contracts, each containing no or limited liability. The money transmitter provides database software to the end-user that typically has limited or no warranties, and the money transmitter carries limited or no liability for providing the software or access. The ISP typically leases a number of telephone lines or telecommunications resources at a certain rate. The underlying service contract with the telecommunications provider is solely for leased space on the network. The network provider, typically one of the large public switched companies, provides only a transport mechanism. This arrangement is similar to right-of-way agreements for utilities or trains that allow use along the track but do not include access to the track. The ISP contracts with the money transmitter for cost-plus as a transport mechanism only, again incurring limited or no liability for this service.

The ISP may enter into a service-level agreement (SLA) with the user (i.e., the money transmitter). Industry standard norms require that the telecommunications system be operational at least 99.5 percent of the time during the service contract. The contract contains a formula for determining an appropriate refund mechanism dependent on the number of times/amount of time access falls below the service level. The money transmitter in turn assumes no liability, or limited liability, to the user. The money transmitter provides no additional value in the form of security for its service; it simply provides a type of bundled service to the user. In essence, the money transmitter charges a convenience fee. The user simply uses the money transmitter's software to create and store the payment order data, which it then sends on a periodic basis to a clearinghouse for deposit or credit to the user's account after it has wound its way through the payments system.

Money transmitters and ISPs that provide services to the financial sector should be required by regulation or legislation to provide liability. Sharing risk is a proven model in the financial services arena, and there is as yet no evidence that this would increase the basic service cost. In fact, only when service entities are required to report losses or suspected losses can sufficient information be garnered to improve pricing for e-security bonds and e-commerce liability insurance.

Lack of a Well-Organized Regulatory Framework for Money Transmitters

Until January 2002, money transmitters in the United States were not regulated at the federal level. However, they are coming under increased scrutiny, because there are now an estimated 200,000 money transmitters operating in the United States and the evidence is mounting that some are being used to launder money. In its 1998-99 annual report, the FATF noted a growing trend to use nonfinancial professional service providers as conduits for money laundering and other nefarious activities. Box 4 outlines how money-laundering concerns have triggered the need to regulate money transmitters in the United States.

As a result of the lack of standardization in regulation and oversight, many money transmitters insert significant risk into the payments system. Typically, they are undercapitalized, use little or no risk-management analysis, and are extremely susceptible to bankruptcy and failure. With the escalation of Internet-related commercial activities and the requisite need to provide ubiquitous payment system conduits, money transmitters are increasing the disintermediation of the traditional payments systems and have a higher profile in the eyes of law enforcement.

Box 4. Money Transmitters and Electronic Security

In 1996, seven money transmitters located on the east coast of the United States were indicted for accepting funds that were allegedly drug proceeds. The El Dorado Task Force, formed in 1992, is a joint federal, state, and local effort involving such entities as the Internal Revenue Service, the New York Police Department, and the New York State Banking Department. It targets industries that facilitate money laundering. That same year, the task force initiated a geographic targeting order (GTO) against 22 money transmitters in New York City and 3,500 licensed agents. The GTO required these transmitters and their agents to report information about cash transfers to Colombia greater than $750. As a result, the volume of money being transferred to Colombia dropped by 30 percent. Under federal law, the government can require a group of financial institutions within a limited geographic range to comply with special record-keeping requirements on a showing of need. Arguably, using a GTO, appropriate federal authorities could require financial services providers to report electronic losses for the same reasons.

Because the primary focus of legislative initiatives targeting money transmitters has been to deter money laundering, most of the activity affecting this industry is derived from anti–money laundering sources.[38]

Two efforts stand out:

1. **The Uniform Money Services Act**, adopted by the NCCUSL in 2000 and known as the Money Transmitters Act.[39] The act requires a money transmitter to obtain a license to operate; sets forth certain licensing criteria, enforcement, and compliance provisions; makes a statement on jurisdiction; and includes provisions on the scope of the act and audit and examination authority. It also contains bond provisions, minimum net worth criteria, provisions on management experience, and requirements that the money transmitter disclose prior litigation and criminal prosecution of management. Only seven states have adopted the act.

[38] See Section V for additional information on money laundering.

[39] See www.law.upenn.edu/bll/ulc/moneyserv/UMSA2001Final.htm

2. **The MRTA Act**, created by the Money Transmitters Regulators Association (MRTA), formed in 1989 as a state regulators organization. Though not as comprehensive as NCCUSL's Money Transmitters Act, it is still a model for dealing with the licensing and regulation of money transmitters. Only five states have adopted it.

Because so few states have adopted these acts, the United States is left with an inconsistent, tedious, and inadequate regulatory scheme. Nevertheless, those states that have shown foresight and initiative in adopting these laws should be able to collect badly needed information on this industry and provide a nucleus from which better regulation can emerge. More exploration is needed to locate the various money transmission channels and regulatory approaches other countries have used. When this paper went to press, none had been located, indicating that emerging markets are not treating these issues systematically.

The last and most promising regulatory effort is enactment of the Gramm-Leach-Bliley Act of 1999. This act affects the future definition of financial services in the United States in the following three ways:

First, the Federal Reserve Board (the Fed) is required to determine what is "financial in nature," taking into account the purposes of both this act and the Bank Holding Company Act; changes in the market and in technology; and an assessment of whether any new activity is necessary or appropriate to compete, to deliver services efficiently, and to offer customers new means of obtaining services.

Second, the Fed is required to decide whether, and to what extent, the following activities are financial in nature or are incidental to a financial activity:

- Lending, exchanging, transferring, investing for others, or safeguarding financial assets other than money and securities.
- Providing devices or means for transferring money or other financial assets.
- Arranging, effecting, and facilitating financial transactions for the account of others.

Third, the Fed may determine that an activity is complementary to a financial activity and by order or regulation deem that activity to be permissible for a financial services holding company.

The Fed has not initiated rules in any of the required or permissive areas. Nevertheless, this act has positioned the Fed to guide the expansion of regulation to include money transmitters and ISPs or any other entity that enables financial institutions to provide services. Thus, the opportunity and the need now exist to initiate global financial forums that call for harmonized approaches to these and other issues raised by the presence of the new market.

Security for Services Provided

ISPs and money transmitters do not necessarily provide additional security for their services. If either is able to offer security, the provider will distinguish between secure and unsecured services. A money transmitter called SWIFT, for example, is careful to distinguish that it provides secure EDI service only. Until a few years ago, SWIFT was a closed system. Today, it has access points to the public switched network. It continues to be one of the most secure transport mechanisms available in the global payments system. FEDWIRE is another example of a closed system, but it now is also connected to the Internet and is subject to vulnerability.

Lacking sufficient terms and conditions in the contract, a user has no way of knowing whether or to what extent an ISP or money transmitter provides security.

Great Britain passed legislation in 2000 that allows the government to track e-mails and seize encrypted Internet communications. It enables law enforcement authorities to demand records of Internet traffic and to view the content of encrypted messages. ISPs are required to set up secure channels to connect to the Government Technical Assistance Center. In turn, the government contributed $30 million to ISPs to cover the cost of installing the "black box" link to the M15 Technical Assistance Spy Center.

Actions to Improve Electronic Security of Payment Systems

The most important objective in a convergent technology environment is to mitigate risk to the extent possible in using an open, universal access architecture. This places greater emphasis on identifying and analyzing systemic risks and vulnerabilities, eliminating risks where feasible, and continually monitoring both risks and security. Few emerging markets appear to have dealt with these issues explicitly thus far. This poses the question of how to do more with less but still increase security and privacy.

In reality, the payment system has broadened and deepened, becoming far more porous and vulnerable. A system is only as secure as its weakest link. Therefore, the first defense recommendation is to enact legislation regulating all money transmitters and any ISPs that provide service to the financial services sector, requiring them to be secure. The Uniform Money Services Business Act would be a good basis for regulating these providers.

Another approach would be to use a request for proposal (RFP) process to shop for value and negotiate the needed terms and conditions in selecting providers. It is important to build in a service-level agreement with appropriate refund mechanisms, liability, and warranties to the terms and conditions.

At present, signing onto the Internet via an ISP results in an adhesion contract in which the vendor dictates all terms and conditions. The industry refers to such contracts as "User Agreements" or "Access Agreements." The contracts are posted on the Internet, and one either accepts the terms and conditions as set forth or does not use the service. Typically, such contracts require the user to check the Internet site periodically for any contract changes, and continued use of the service constitutes acceptance of the terms and conditions. Adhesion contracts, once considered unenforceable, are becoming the norm in the ISP and electronic-commerce-dominated industries, especially the financial services industries.

Another avenue of defense is self-regulation through the automated clearinghouse process or, more broadly, via specific arrangements outlining security standards in the case of wholesale or retail payment networks. Building clearinghouse rules requiring all entities to use vendors that provide an appropriate level of security and to post sufficient money or bond to cover losses would create an incentive for the parties to establish a proper electronic security standard. This approach needs to figure more prominently in the ongoing work of establishing wholesale and retail payment networks in emerging markets. Moreover, as in the case of securities regulation, central bank supervision of SROs that are responsible for retail or wholesale payments will become far more important.

Insurance coverage is yet another means of protection. Financial services entities should use insurance to protect themselves from gap loss, whereby e-risk is realized even after insurance companies have required a financial services provider to meet specific security standards. Section VIII will examine this issue in more detail.

VII. Pillar III: Supervision and Prevention Challenges

Background: Electronic Security and E-Banking Supervision

In 1999, the Basel Committee established the Electronic Banking Group (EBG) to focus on adapting the Basel Committee Guidance as necessary to e-banking issues. Moreover, the Financial Stability Forum (FSF) has established a special overall contact group that is in the process of discussing what issues need to be addressed in the implementation of the 14 principles identified by the EBG (see Box 5).

Box 5. Principles for Managing Risk in Online Banking

The Electronic Banking Group of the Basel Committee on Banking Supervision has identified 14 key risk-management principles for online banking. Banks and their supervisors should consider these principles when formulating risk-management policies and processes for online activities.

* Management oversight. Effective management oversight of the risks associated with e-banking needs to be in place, and e-banking risk management should be integrated with overall risk management.

* Management of outsourcing and third-party dependencies. Comprehensive, well-defined, ongoing oversight is needed for managing outsourced relationships and third-party dependencies supporting e-banking, including adequate prior due diligence.

* Segregation of duties. Appropriate measures are needed to ensure proper segregation of duties in e-banking systems, databases, and applications.

* Proper authorization measures and controls in systems, databases, and applications. Appropriate authorization measures and proper controls need to be in place for e-banking systems, databases, and applications.

* Clear audit trail for e-banking transactions. A clear audit trail is needed for all e-banking transactions.

* Authentication of all entities, counterparts, and data. Banks should authenticate the identity and origin of all entities, counterparts, and data transmitted over the Internet.

* Nonrepudiation (accountability) for e-banking transactions. Nonrepudiation should be ensured to hold users accountable for e-banking transactions and information.

* Comprehensive security control. Banks should ensure the appropriate use of activities and properly safeguard the security of e-banking assets and information.

* Integrity of transactions, records, and information. Banks should prevent unauthorized changes to and ensure the reliability, accuracy, and completeness of e-banking transactions, records, and information.

* Appropriate disclosure. To avoid legal and reputation risks, including risks for cross-border activities, banks should have adequate disclosure for e-banking services.

* Confidentiality and privacy of customer information. The confidentiality of customer information and adherence to customer privacy requirements should be ensured.

* Business continuity and contingency plans to ensure the availability of systems and services. Plans should ensure that e-banking systems and services are available to customers, internal users, and outsourced service providers when needed.

* Incident response planning. Incident response plans should be in place to manage and minimize problems arising from unexpected events—including internal and external attacks that hamper the provision of e-banking systems and services.

* Role of supervisors. Bank supervisors should assess banks' management structures, practices, internal controls, and contingency plans for e-banking.

Source: Electronic Banking Group of the Basel Committee on Banking Supervision; see also Annexes I–III of this paper to understand how such principles can be translated to distinct processes to reduce e-security risk.

Because e-banking is based on technology designed to expand the “virtual” geographic reach of banks and customers without necessarily requiring a physical expansion, market expansion beyond national borders significantly increases cross-border supervision challenges for bank supervisors. Although such supervisors agree that the supervisory principles of traditional banking are applicable to e-banking, changes in technology and dependence by banks on service providers magnify the level of risk. The 14 principles for risk management of e-banking issued by the EBG fall into three fundamental categories: (1) effective board and management oversight, (2) security risk issues, and (3) reputation risk issues.

The ability of regulatory agencies to regulate and supervise e-banking entities effectively in today’s virtual banking environment must be strengthened to handle the special challenges of electronic security. Authentication, security control, integrity, and even incident response planning figure prominently in the 14 EBG principles. In particular, the EBG emphasizes the need for a bank’s effective internal controls. Moreover, the EBG principles place liability on the banks in the event of electronic security problems with vendors. Despite this emphasis, there is still a need to make the chain of vendors involved in the delivery of electronic security services or other e-enabling services secure and to impose better downstream liability on these entities. For example, the Office of the Comptroller of the Currency in the United States has done extensive work to draft “electronic security” guidance for U.S. banks and vendors.

In many countries, a bank is subject to examination on a periodic basis. In the past, traditional examinations were done on-site and based on safety and soundness through the CAMEL rating system.[40] In addition, banks in most countries throughout the world are subject to some variant (where weights may differ) of the Basel capital adequacy guidelines. The challenges presented by electronic security breaches are not explicitly accounted for in this framework and, as noted below, even the present capital standards do not really explicitly address this form of risk.

Bank Capital Standards and E-Security

In May 2001, the Basel Committee on Banking Supervision issued a consultative document relating to capital adequacy regulations. This document defines operational risk as the “risk of direct or indirect loss resulting from inadequate or failed internal processes, people, systems, and external events.”[41] It identifies three ways to measure operational risk: (1) the basic indicator approach, (2) the standardized approach, and (3) the internal management approach. Under the basic indicator approach, banks have to hold capital for operational risk that is equal to a fixed percentage of gross income. In the case of the standardized approach, a more complex process is used whereby the financial services provider breaks up its overall operations into distinct business lines and uses different indicators for each and then computes the capital charge via use of a capital factor provided by supervisors. Finally, the most advanced approach is the internal measurement approach, which relies on calculations that result in expected losses.

None of these frameworks allows for what one might think of as kidnapping- or extortion-related risks caused by penetration of a bank’s systems. Moreover, the concept of operational risk that is now used addresses only legal risk, not the problems of strategic and reputation risks. Since incentives to report losses or compromises of the system accurately are

[40] Capital Assets Management Equity and Liquidity (CAMEL) is a system that is based on a ranking of one to five, with one being the best.

[41] See Basel Committee on Banking Supervision Consultative Document: The New Basel Accord**,** January 2001.

often lacking, taking proper account of electronic security risks in any concept of operational risk will be highly subjective and complex.

E-Security and IT Examination Processes

What, then, is the best way forward if capital regulations cannot be adjusted? One of the most fruitful avenues is to publicize the actions that can be taken to measure and manage the risk of electronic security breaches and for regulators to be pro-active in providing guidance as well as modernizing the approaches to on-and off-sight supervision of such electronic transactions as is occurring in some countries. Implementing new guidelines and risk-management processes that can be monitored by bank examiners would impose a minimum standard for dealing with electronic security because it could reduce the prospect of security breaches. Here, adoption of some form of layered electronic security risk protocol might also be worthy of consideration. Box 6, which draws on extensive consultations with electronic security industry experts, illustrates such a set of layered security measures (see also Annex I, which contains more detail). A bank could have many of these layers of security in place. A number of these actions are not costly to implement for any financial services provider, yet they are often lacking.

In recent years, IT examinations have been performed on banks that possess online transactional banking systems. IT examiners would often enter a bank and ask the following questions:

1. Do you have a firewall?
2. If so, is it configured properly?
3. Do you possess a local area network (LAN) or wide area network (WAN)? If so, are there encrypted channels?

Recently, a number of countries, including the United States, have passed legislation stipulating the need for financial services providers to strengthen their information security. For example, the GLBA, also known as the Financial Services Modernization Act or Title V 12 CFR 573, applies to "financial institutions." These are defined very broadly in Section 509(3) of the act to mean "any institution the business of which is engaging in financial activities described in section 4(k) of the Bank Holding Act of 1956." GLBA states that these institutions must adhere to the following actions:

- Identify and assess the risks that may threaten customer information.
- Develop a written plan containing policies and procedures to manage and control these risks.
- Implement and test the plan.
- Adjust the plan on a continuing basis to account for changes in technology, the sensitivity of customer information, and internal or external threats to information security.

Essentially, GLBA addressed the pivotal question, “What is being done to secure customer data, both physical and electronic in origin?” As in the case of payment system security this is a step in the right direction. However, the law needs improvement vis-à-vis the specifics as to how banks should protect their electronic information assets. The 1996 Federal Financial Institutions Examination Council's (FFIEC) IT examination manual has been the industry norm, and it is currently undergoing an important update. More specifically, the Office of the Comptroller of the Currency (OCC) in the United States had issued extensive “guidance” on security; and, in partnership with other regulators in the FFIEC, a new IT security booklet for banks is under preparation.[42]

Box 6. Layered Security

In today's business climate, the gap between risk management of physical assets and informational assets is large. Layered security is composed of 12 core elements. Annex I presents a toolkit for effective risk management compiled from the contributions of industry leaders and law enforcement officials. The following categories represent the 12 core layers of e-security where security is a dynamic process, and, therefore, such policies and processes must constantly be reviewed.

Information Security Officer—The creation of the position of Chief Security Officer who overseas that the other 11 layers are carried out and implemented in accordance with the best practices laid out in Annexes 1 and 2.

Risk Management—A broad based framework based upon CERT's OCTAVE paradigm for managing assets and relevant risks to those assets.

Access Controls/Authentication—Establish the legitimacy of a node or user before allowing access to requested information. During the process, the user enters a name or account number (identification) and password (authentication). The first line of defense is access controls; these can be divided into passwords, tokens, biometrics, and public key infrastructure (PKI).

Firewalls—Create a system or combination of systems that enforces a boundary between two or more networks. Annex I contains recommendations for proper firewall configuration.

Active content filtering—At the browser level, it is prudent to filter all material that is not appropriate for the workplace or that is contrary to established workplace policies.

Intrusion detection system (IDS)—This is a system dedicated to the detection of break-ins or break-in attempts, either manually or via software expert systems that operate on logs or other information available on the network. Approaches to monitoring vary widely, depending on the types of attacks that the system is expected to defend against, the origins of the attacks, the types of assets, and the level of concern for various types of threats.

Virus scanners—Worms, Trojans, and viruses are methods for deploying an attack. A virus is a program that can replicate itself by infecting other programs on the same system with copies of itself. Trojans do not replicate or attach themselves to other files. Virus scanners hunt malicious codes. Annex I details proper maintenance and configuration of these scanners.

Encryption—Encryption algorithms are used to protect information while it is in transit or whenever it is exposed to theft of the storage device (e.g. removable backup media or notebook computer).

Penetration testing—Penetration testing entails obtaining knowledge of vulnerabilities that exist on a computer system or network and using that knowledge to gain access to resources on the computer or network while bypassing normal authentication barriers.

Proper systems administration—This should be complete with a list of administrative failures that typically exist within financial institutions and corporations and a list of best practices.

Policy Management Software—Bank policy and procedural guidelines vis-à-vis employee computer usage should be controlled by a software program.

Business Continuity/ Incident response plan (IRP)—This is the primary document used by a corporation to define how it will identify, respond to, correct, and recover from a computer security incident. The main necessity is to have an IRP and to test it periodically.

All employees should be aware of the correct procedures in the event of a computer incident. See Annex I for more detailed treatment of these issues.

Until recently, IT examiners had followed guidelines that were, in effect, a modified version of the old FFIEC IT examination manual. These IT examiners perform "risk scoping," a practice wherein they only check new systems or software installations that have occurred since the last examination. If the examiner has checked an institution in the past and given it a good

[42] Please see OCC Bulletins: 2001-35, 2001-12, 2001-8, 2000-14. See also various Alerts and Advisory letters.

score, he or she will not recheck any of the older systems and configurations. This approach, however, can be highly problematic. Systems change, and new vulnerabilities in software and configuration appear daily. Examiners now need to assume that systems checked in earlier audits may no longer be secure.[43] Today it is no longer appropriate to react, government is expected to be proactive to anticipate, plan and be prepared. If the practice of risk scoping exists merely to save time and costs, legislatures should mandate additional funding for regulatory agencies.

Hosting companies such as FiServe are examined by joint examination teams from the Office of the Comptroller of the Currency (OCC), the Federal Reserve, Federal Technology Services (FTS), and the Federal Deposit Insurance Corporation (FDIC). The Bank Service Corporation Act states that if an entity provides a data processing service to a bank, then it, too, can be examined. These entities, however, cannot fail the exams. The examiners note deficiencies, and then the entity and examiners agree to a plan of action. If negotiation fails, the enforcement action calls for implementation of a cease-and-desist order. Yet again there is a loophole. Because no real reporting requirements are in place for these hosting providers for losses or rates of intrusions, the cease-and-desist "stick" is negated because there is no information on which to base it. Hence, no standard exists for the evaluation and subsequent regulation of e-security in banking institutions.

Supervision will need to be proactive, given the hostile nature of the Internet environment. As far back as 1995, the ISO/IEC 13335, better known as the Guidelines for the Management of IT Security (GMITS), recognized that the Internet was a hostile environment that would require the use of proper electronic security.[33] Box 7 outlines the processes that were advocated. Note that the layered electronic security risk analysis advocated in this paper (see Annex I and Box 6) has many similarities to this ISO standard, which has not been well implemented in many types of institutions, including banks.

Toward a New Approach to Regulation and Supervision

Redefining Regulatory Authority and Legal Liability of Downstream Vendors. Regulatory agencies need improved powers and the appropriate authority to regulate fully all third-party money transmitters and ISPs. Their budgets and legislative tools will need to increase and the means found to rely on auditing companies (if properly reformed) and the insurance sectors of emerging markets to play a role in this process. The following regulatory and compliance actions might help mitigate the threat of system compromise yet not overzealously extend the safety net. In addition, adoption of processes to monitor the extent to which financial services providers adopt and employ better layered electronic security risk-management practices will be essential as part of any enhanced regulatory and compliance regime.

[43] Nor should they assume that the snapshot provided by onsite examinations is an accurate picture of real time activity.

Box 7. ISO/IEC 13335 Information Technology—Security Techniques—Guidelines for the Management of IT Security (GMITS)

This ISO/IEC technical report, published in 1995, is generally known by the acronym GMITS. It is made up of five parts, designed to address different aspects of Internet security.

Part 1. Concepts and Models for IT Security
Part 2. Managing and Planning IT Security
Part 3. Techniques for the Management of IT Security
Part 4. Selection of Safeguards
Part 5. Management Guidance on Network Security

GMITS was written to be usable and useful in the worst-case environment; that is, a hostile environment, such as the Internet. The properties of assets (information) that need to be taken into account and protected are extended from the classical confidentiality, integrity, and availability to include accountability, authenticity, and reliability.

Vulnerability is refined to include any property of the asset that can be exploited for purposes other than intended. Thus, a firewall represents a single point of failure and is susceptible to a denial-of-service attack, which does not detract from its value as a protections mechanism but does mean that this vulnerability needs to be considered and addressed.

Likelihood is refined to be associated with use of the data to perform risk analysis, risk assessment, and risk management.

Part 3 focuses on the topics of risk analysis, risk assessment, and risk management. GMITS recommends that the organization establish a baseline minimum set of controls that will be applied to all aspects of the organization. This baseline will be maintained through the use of a median level risk analysis. Policy is used to ensure the enforcement of the baseline throughout the organization, so that all areas can rely on it.

There are never sufficient funds to implement the ideal set of safeguards, and thus safeguards that provide multiple functions are to be preferred, provided the compromise does not reduce effectiveness. The most important situation to guard against is a false sense of security, which is actually worse than having less security or no security at all. A modicum of paranoia is a good thing.

Having been developed as a generalized document, GMITS does not address in detail particular aspects or subtopics of IT security, such as network, cryptographic, or emanations security.

Regulatory

- Expand the circle of regulated entities to include those elements that traffic in or assist in money transmission and directly connect to any payment system.
- Review regulatory goals and needs in an electronic environment.
- Train audit and examination special teams in risk analysis, risk management, and IT issues.
- Revisit capital adequacy requirements and the definition of operational risk to evaluate how best to accommodate e-risks noted in this paper.
- Provide report cards to the public on how well the financial services industry is doing to attain the new security objectives in this area.
- Require clearer management responsibility and accountability to create and sustain safety and soundness.
- Define the regulatory paradigm for the new market.

Compliance

- Develop analytical teams to assess and monitor e-risk management.
- Disconnect any entity from the system that is not in compliance.
- Require warranties, indemnification, and liability from service providers that connect to the payments system.
- Require insurance coverage to accommodate additional risk.
- Institute well-developed reporting requirements for all electronic money or electronic data losses from all service providers and financial services entities.
- Require information sharing between the regulator and the financial services entity concerning losses.
- Require artificial intelligence software, and make affirmative the duty to report all irregular activity from or through any service provider.
- Ensure that in management letters and other correspondence between examiners and management of financial services providers adequate attention focuses on communication between the systems administrator and senior management and even the board of directors.[44]

Access, availability, and interoperability should also be key objectives of supervision and enforcement. The very interlinked nature of electronic security providers and e-enabling companies or money transmitters implies that the traditional regulatory structure must expand. It does not imply that a greater number of entities be under the safety net but rather that the regulatory framework create incentives for accountability in such entities as ISPs to application service, software, hardware, monitoring detection, and assessment providers. Liability must attach to these providers just as to the directors of those financial institutions that contract with them. These providers are as indispensable to the institution's ability to provide electronic financial services as lawyers and accountants. They should be bonded, licensed, and subject to periodic audit and examination.

In sum, traditional regulatory schemes are outdated and cannot adequately address the new components of the payment system to determine whether a financial institution is operating in a safe and sound manner.

Coordination in Supervision and Information Sharing Across Agencies

In many countries throughout the world, supervision and enforcement in the area of electronic security is complicated by unclear jurisdictional lines across relevant agencies. In practice, often the central bank, the securities or banking regulator (if separate from the central bank), law enforcement agencies, and many other entities must be in a position to share information and reports. In many cases, this can be problematic from a legal point of view, or a general lack of incentives may result in no established forum or process for undertaking coordinated action.

It is important to seek and promote cooperation between law enforcement agencies and regulatory authorities for financial services providers. Increasingly, such cooperation will be needed within and even across countries. Such arrangements will have to go beyond the pursuit of

[44] During the Y2K effort, systems administrators were given more attention, but in many financial services conglomerates, very little communication goes on between management and the systems people until after the fact. As technology budgets and related security issues grow in importance, this is likely to change—but the regulatory authorities can make management more sensitive to these issues in the course of the examination process.

those engaged in money laundering activities; it will require the development of a more accurate and timely system for reporting all incidents of electronic security breaches, and not just loss-related information. This is an important area, in which worldwide cooperation will be needed on an increasing scale.[45] To achieve such cooperation may require greater harmonization across countries in fundamental areas of legislation, including bank secrecy statutes.

VIII. Pillar IV: The Role of Private Insurance as a Complementary Monitoring Mechanism[46]

Background

Despite formidable reportage problems inherent in establishing a benchmark to actuarially measure the risk of hack attacks, electronic identity theft, and other forms of related e-risk, insurance companies are writing coverage for such risk. The development of e-risk policies first occurred in the mid-1990s when insurers recognized the coverage gaps or gray areas in traditional insurance products for perils on the Internet. In response to those risks, insurers developed stand-alone e-risk policies rather than adding coverage to existing property and liability insurance. Market participants have used employee liability coverage as a model for pricing and issuing this insurance.

In underwriting this risk, insurers combined information security standards, such as the BS7799, with principles of risk management that included analysis, avoidance, control, and risk transfer. Today, insurers recognize the ISO 17799 information security standard, which addresses these issues in the following 10 major sections:

1. Business continuity planning
2. System access control
3. System development and maintenance
4. Physical and environmental security
5. Statutory, regulatory, or contractual obligation compliance
6. Personnel security
7. Security management for third-party access or outsourcing to a third-party service provider
8. Computer and network management to safeguard information assets
9. Asset classification and control
10. Security policy management support

As part of the e-risk application process, several major insurers, including AIG, Zurich, Chubb, St. Paul, Progressive, and Lloyd's, have incorporated the ISO 17799 standards into a baseline security questionnaire that becomes part of the insurance application in e-risk policies they underwrite. In order to bind coverage, the insured must meet a certain security threshold for insurability, and the precise nature of such thresholds has not been completely standardized within and across countries. In part, this reflects the very dynamic impact of technology in this area. Despite these developments, the use of e-risk policies is still nascent.

[45] See Section X, which includes a few examples of such cooperative ventures as Computer Emergency Response Teams (CERTS) or the New York Electronic Crimes Task Force.

[46] The authors thank Kurt Susse of Galaxy Computing International for very helpful written contributions to this section.

In the case of first-party coverage, such policies are being explicitly designed to provide coverage against network extortion, computer theft, damage to digital assets and information as intellectual property, and business or dependent business losses. In the case of third-party coverage, such policies are designed to cover network security or loss event liability and electronic publishing and multimedia liability.

In underwriting these special e-risk policies, insurers are increasingly assessing the extent to which specific providers of financial or other services are in compliance with appropriate standards in each of the 10 areas specified under ISO 17799. These areas are also relevant in the design of appropriate layered security systems, such as the guidelines in Annex I of this report. These types of considerations still do not make it possible to actuarially calculate proper premiums for these forms of first- and third-party e-risk coverage. The underlying defects in the information about intrusions and extortion make the pricing of such policies anything but straightforward.

ISO 17799's Failures[47]

There are six inherent weaknesses within the approach laid out by ISO 17799 standard. The first relates to pricing. Due to lack of significant loss history and unknown potential for loss, carriers are proceeding on cautious side and making premiums cost prohibitive. This presents an obvious quandary. The second relates to underwriting. There exist an over-dependence on third party assessments to qualify risk and lack of qualified in-house staff. Banks and corporations are not rewarded for their pro-activity in security measures and are lumped into "hazard classes". The third such weakness relates to a crisis in communication. If a carrier does require a third party assessment; it is done once a year and is "static". Due to rapid changes in security, there is great potential for carriers to be insuring companies with significantly different profiles than original assessment. The fourth weakness corresponds to the loss of control by banks and corporations. These entities are paying substantial premiums and receiving a negligible amount of education on new vulnerabilities and subsequent proactive risk exposure procedures. The fifth weakness corresponds to the cost and severity of claims as a result of security breaches. Due to lack of case law and precedence, it is assumed to all security breach losses will be "limited losses". This is not necessarily true. Especially since organized hack attacks are prevalent in today's world wide web. Finally, the most critical of all of ISO 17799 failures is the static nature of policies. The rigidity of policies and the subsequent inability of consumers and underwriters to edit policies to specific to risks / legislation associated to a specific enterprise or bank can undermine the coverage entirely.

In thinking about the future coverage for this standard, one must consider:

1. How identified vulnerabilities will be linked to potential subsequent losses and applicable coverage grants.
2. How uniformity of vulnerabilities under review by auditors will be addressed.
3. How financial industry specific vulnerabilities will be assessed.

Traditional Insurance Policies

Typical insurance policies have not dealt with electronic security risks or, more broadly, the types of risks emanating from such security breaches. For example, so-called first-party coverage in the context of commercial property policies usually requires physical loss or damage

[47] Contributed by Tim Burke, Director of Business Development, Riskology Inc. www.riskology.net

to property via fire but not denial-of-service attacks via computer hackers or other types of e-risk. Also, an employee theft exclusion is usually included in such policies; in many cases of electronic security breaches, an insider or former employee may be involved. In fact, in Fall 2001, the insurance service office explicitly excluded software- and computer-related losses in commercial policies so that coverage would need to be sought via other specialized policies or arrangements. Commercial and crime policies generally cover theft of money and securities, not theft of information, as do many forms of fidelity bonds. Finally, kidnap and ransom policies often limit coverage for extortion to threat of bodily injury, not to the possibility of severe reputation damage associated with making public penetration into a bank's systems or theft of other information.

Recently, insurance carriers have been offering e-risk policies that do provide cover against cyber risk. Here there is the broader question of how to characterize the specific risks to reputation entailed in electronic security breaches and—because reputation risk is highly complex—the kinds of loss payouts for which insurance carriers would be liable. One could just as easily view these risks as similar to catastrophic risk, or perhaps even to kidnapping risk. The latter is relevant not only in the case of electronic identity theft, in which a ransom may be sought from the financial services provider, but also in the case of a pure hack where the hacker threatens to go public and may demand what amounts to a form of extortion payments. Defining the nature of the risk in the case of first-party coverage deserves more thought in light of how industry participants are now writing such e-risk policies.

Another form of insurance that is generally not adequate is third-party coverage. Here there have been gaps in the narrow provisions for advertising injury coverage in which claims can be sought only if the injury occurs in the coverage territory during the policy period—thereby excluding many possible electronic security events. Despite refinements made to the definition of advertising on the Internet via the electronic data liability amendment in Fall 2001, this is an area that remains unresolved. Also, because electronic data is not defined as tangible property, these forms of coverage have limited effectiveness.

Finally, many of the actual e-risk policies reviewed in preparing this report pay no attention to the special risks that wireless technologies are creating in the delivery of financial services. As documented in Annex III below and in a separate paper, Mobile Risk Management, insurance providers should clearly identify the standards for financial services providers to meet for wireless risk mitigation before they underwrite an e-risk policy. In so doing, the insurance industry could play a critical role in setting standards for electronic security risk mitigation.

Insurance Companies as a Force for Change

Over time, the growth in e-commerce liability insurance and, specifically, e-risk insurance is likely to be quite substantial. Estimates by AIG suggest that the market for this insurance may be as much as $2.5 billion.

The viability of providing this insurance coverage is related to more systemic approaches to improving the base of information for pricing[48] the electronic security risks to be covered. Although vendors of electronic security services are working with insurance companies on this issue, government, industry, and law enforcement officials clearly need to find ways of improving

[48] QUAID (Questions Used to Access the Information Database) is an active database storing past, present & ongoing assessments along with known vulnerabilities dating back to 1985. Specifically developed for the transfer of risk through the creation of new pricing models for E-risk. www.riskology.net .

the reporting of such information (see Section IX). Current efforts to develop public-private partnerships to solve this problem should therefore be a high priority.

The global insurance industry can and should act as an important force for change in electronic security arrangements worldwide. First, it should strive to improve the minimum standards for electronic security and should strongly advocate enhanced layered electronic security systems (see Annex I). Second, it will be interested in improved certification standards for vendors of electronic security services described in Section III as a way of mitigating risks of coverage and of spreading risk. Third, it will be concerned with improvements in worldwide cooperation and efforts to improve the data and information available with which to actuarially measure e-risks in companies and financial services providers. Finally, it will favor solutions that require vendors of electronic security and other related services (e.g., hosting) to bear some liability, in contrast to some of the current arrangements, which are entered into by parties in the financial services industry in outsourcing arrangements and do not create adequate incentives to maintain electronic security.

IX. Pillar V: Certification, Standards, and the Roles of the Public and Private Sectors

Four potential areas of certification to address in the electronic environment are the following: software, hardware, IT security vendors, and electronic transactions. Software and hardware vendors were discussed earlier in the paper. Here the main concerns are that hardware and software vendors often provide products with known vulnerabilities that should not be used for financial transactions. Yet they sell these products and refuse to provide warranties or liabilities for them. The industry could provide certifications for these products, but a better approach would be to require vendors to warrant their products and provide either liability coverage or notice and disclaimers when a product is not suitable for certain uses.

Next are questions about the roles of government and the private sector in certifying aspects of electronic financial services, and the issue is broader than just how it relates to the PKI. First is the question of whether there is a case for regulators to license vendors that provide electronic-security-related services to the financial sector. Such vendors play a role in protecting the integrity of one of the eight critical infrastructure components of the electronic economy. However, licensing vendors would widen the regulatory safety net. Might another alternative provide assurance without unduly burdening the regulatory structure? For example, such vendors might post a form of performance bond, or they could be required to obtain professional liability insurance through private insurers. Or the industry could require them to obtain certification levels, enabling them to provide certain services based on the level of certification achieved.

Probably, industry regulation through a certification process will yield the most consistent results, particularly if insurance provides incentives to certified vendors as well as to institutions that use such vendors. This way, regulators can require vendors to share in the risk through professional liability. Only those parties essential to the delivery of the financial services would be included in the regulatory net, security would be a prerequisite for providing services to the financial sector, and all would share proportionately in the attendant risks. Thus, the scope of regulation could be contained to those entities, such as money transmitters and ISPs, that hold themselves out as being able to provide hosting to the financial services industry. The steps in brief are for industry to certify vendors to levels of professional ability, have insurance concur through coverage or performance bonds, and have risk appropriately shared.

At the transactional level, as part of its business practice, an institution should analyze the benefits that each technology solution brings to the table and weigh that against the costs or concerns associated with each. Then it should implement a data security classification system through the business rules engine mechanism that automatically attaches a level of security to each type of transaction. The business criteria used to make these decisions should include at a minimum the following value matrix: integrity, reliability, authentication, verification, authority, and nonrepudiation. The value of a transaction should then be equal to the sum of the total risks associated with the transaction.

Using such a value matrix could also assist the insurance industry in evaluating coverage risks and pricing. Moreover, it could help the financial entity with self-monitoring by pinpointing where and why particular risks are greater. The value matrix would also help to enrich the information that is reported. The institution could use a mix of solutions, fitting the solution to the value and risks of the underlying transaction. Although insurance companies could play a role in encouraging the security industry to set standards and even to endorse best practices in terms of authorizing and verifying transaction elements, setting harmonized standards for authenticating documents and such related issues goes beyond the role of any private entity and requires significant cooperation between governments.

Traditionally, encryption has been used as a means to protect the information transferred over the Internet, together with various types of protocols (e.g., secure socket layer, fix, and others) designed to provide security to naked or "open" wide area network systems. Although effective, these mechanisms are meant only to provide protection against certain kinds of vulnerabilities.

The process of securely transmitting information over the Internet in countries or across countries has led to a proliferation of public and private key providers and related "certification authorities." These services can be provided by government agencies, such as postal authorities; by technology providers, such as GTE or Verisign; by telecom service providers, such as Nortel's Entrust; and by financial services providers. Eight global financial institutions are such providers.[49]

Every user of public key cryptography is freely provided a key. The creation and storage of such keys, as well as the attendant certification processes, present major challenges. As Annex II shows, there are many ways to authenticate that can be used along with encryption.

First, it is necessary to address the development of a proper certification process for public and private keys and the levels of use of the process. Some countries have opted to endorse only one recognized public certification authority (such as the postal service). In other countries, both public and private authorities provide this function. Although one could claim that certification is a "public good" and therefore should be kept under the control of a public entity, such as the post office, private companies could act as certification agents as long as there is a viable means of cross-certifying to check on the competence of the service being provided. In all likelihood, the desire to maintain the institution's reputation will act as a significant incentive to resolve the moral hazard problem.

[49] The certification authority authenticates the public key by distributing it with a certificate (digitally signed by the certification authority). The potential liability of the certification authority, as well as the reputation implications of security-related breaches, have been used as an argument for the outsourcing of the public key infrastructure to private providers. The seven banks that are certification authorities are ABN, Bank of America, Deustche Bank, Barclays, Chase, Citigroup, and Hypoverensbank.

Second, governments need to address the issues of authentication, confidentiality, and nonrepudiation in designing valid electronic transactions, because these form the backbone of transactional activity. Annex I discusses these issues in detail and compares the benefits and drawbacks of potential technologies, such as biometrics and digital time stamping. More generally, government needs to encourage the development of technologies that can be used to authenticate with or without certifying. To preserve confidentiality, the government can require the double signing of a key or the use of certain encryption. Again, government should encourage the private development of solutions that maintain confidentiality and privacy for businesses and consumers. In fact, a global industry has already developed, and many U.S. companies are providing privacy and security solutions to companies and consumers worldwide, as noted in Section III.[50]

Third, the integration of technologies through multiple channels for delivery of financial services, as noted in Section III, implies the need to explore how best to harmonize standards across countries. Technologies will need to interface, but sufficient security must be in place so that commerce can be conducted across countries even if they have different forms of certification. The ISO standards could play a constructive role (see Box 2 in Section III), but the challenges will not be small.

Finally, it is important to consider how to ensure an appropriate level of trust in any given transaction. The legal or regulatory transactional framework must be technology-neutral. In reality, a variety of technologies can certify or authenticate transactional elements and can protect against nonrepudiation. The next subsection reviews the major technologies in use today and examines their strengths and weaknesses.

Trust and Confidence in Authentication Technologies and Certification. **“**Trust and confidence” translates into the following: Party A is able to access online services and transfers funds from one account to another. Party A then checks his account balances, and the correct amount has moved from one account to the other. At the end of the month, he goes online again and confirms that all activity for that month has been properly posted and that the account balances match his figures. As a result, he has a high level of trust and confidence in the system. Or Party B receives certain monies from the government on a monthly basis. Or Party C sets up automatic bill paying for all her utilities. Each month, her account is debited for the correct amount of the utilities. Studies have shown that when someone uses a new technology, that party will bond with the use of the technology if it works favorably with no complications the first three times of use. Conversely, assume that Party D approaches an ATM and attempts to take money from his account. He inputs his personal identification number, and the transaction is refused. He tries again, and it is refused again. The third time, the ATM machine eats his card. Studies show that the opportunity to create trust in the technology has been lost. This person will not willingly use the technology again unless no other delivery channel is available.

PKI Technology. An extraordinary amount of research and development money has been spent on developing PKI and certification authorities over the past decade. As a result, PKI is the best known electronic signature verification technology. (See Annex II.) Clearly, it has its strengths. But easier and simpler technologies perform just as well. Again, it is important to understand the business drivers and the consequential risks in choosing an appropriate technology. Moreover, there is no accepted standard legislation, and record retention requirements for certification authorities are often undefined.

[50] These solutions include systems providing safety in browsing to detect cookies or manage cookies; e-mail security; and even personal firewalls for retail consumers.

Notaries. One alternative to PKI is to offer a new type of notary license. In this scenario, a notary could apply for a Class A license. This would authorize the notary to accept and certify digital and biometric signatures and to time-stamp documents and notarize manual signatures. Or a notary could apply for a Class B license. This would authorize the notary to time-stamp and notarize manual signatures only. Or the notary could apply for a Class C license. Under this scheme, the notary could only notarize manual signatures. This multi-license notary scenario is a tempting resolution to the issue of nonrepudiation for a number of reasons. First, it simply expands an existing, accepted, and regulated framework for verifying signatures. It assesses a greater fee for a Class A license than for the others, and this in turn acts as a user's fee, which can be used by governments to pay for the necessary personnel and equipment to provide online assistance to users and to the expanded notary industry. The negatives of such a solution are also fairly clear. In emerging markets, notaries may not be well trained to undertake this role, and they would need to receive certifications to perform this function. Another concern is that the licensing system, or in many cases the notaries themselves, may be subject to corruption; this concern emphasizes the need for sufficient oversight. Moreover, in the context of many transactional arrangements, notaries often increase the costs of transactions.

Digital Time Stamps. Another alternative to certification authorities is a digital time stamp (DTS) service provider. A time stamp associates a certain date and time with the creation of a digital document. The time stamp can be referenced to prove that the document was recorded at a specific date and time. For example, Party A signs a document and wants it time-stamped. She computes a message digest of the document using a secure one-way hash function and sends it to a DTS service. In return, the DTS service sends back a digital time-stamp document. This includes the message digest, the date and time it was received by the time-stamping service, and the digital signature of the time-stamping service. Later, Party A presents the document to verify its creation date, and a verifier recomputes the message digest and determines whether it matches the digest in the original time-stamped document. The verifier then verifies the digital signature of the time-stamping service. The strengths of this process are that a message digest does not reveal the contents of the document but simply verifies that the underlying message was received on a certain date and time. As stated, a DTS could be an added dimension to a notary's license. In addition, or separately, the DTS could be provided by the post office for set fees. Again, this would use an existing entity that is familiar to the consumer.

Biometrics and Certification. Biometrics is another alternative to the verification process. Biometric authentication techniques can be used to verify the identity of people online automatically through their distinctive physical or behavioral traits. A biometric identifier represents a physical characteristic of the user (see Annex II). The global recognition of this authentication technology will assist in the nonrepudiation of financial transactions and subsequent documentation. These technologies facilitate the process by which entities can transact on a medium that facilitates anonymity. In this case, the two issues to address would be (1) certifying the specific biometric technology and its accuracy, and (2) defining a digital signature in a broad enough manner to allow certification of the parties to a transaction through whatever authentication technology makes sense.

In summary, government should let the private sector lead where possible but should temper this approach by adopting open standards; endorsing technology-neutral solutions; encouraging the industry to self-regulate and certify; and helping insurance and other industries use incentives to share risk and responsibility in identifying and correcting vulnerabilities. Such objectives can be difficult to achieve in emerging markets where oversight of such self-regulatory entities can be defective.[51]

X. Pillar VI: Accuracy of Information on E-Security Incidents and Public-Private Sector Cooperation

One action that would improve electronic security worldwide would be the creation of a set of national and cross-border incentive arrangements encouraging financial services providers to share accurate information on denial-of-service intrusions, thefts, hacks, and so on. Ample evidence shows, as noted in Section II, that no accurate base of information exists either within or across countries. This situation limits both awareness and the scope of private sector solutions that can be provided and may even be increasing the cost to companies and financial services providers of insuring against such risks.

Prompted by law enforcement, industry participants, and the academic community, greater public-private cooperation is starting to become more of a reality in the United States and, increasingly, in many other countries as well. Some innovative examples of such efforts, but by no means the only ones, are described below.

The Internet Security Alliance (www.isalliance.org) and the Computer Emergency Response Team (CERT).[52] This is a collaborative effort between Carnegie Mellon University's CERT Coordination Center and a cross-section of private international companies that include NASDAQ and Mellon Financial, TRW, and AIG. This alliance is an industry-led, global, cross-sector network focused on advancing the security of the Internet. CERT (see Glossary for detail) is expanding its operations and now has counterparts in more than 140 countries. It is beginning to implement its methods for extracting this information from users on a global basis.

The Forum of Incident Response and Security Teams (FIRST). FIRST brings together a variety of computer security incident response teams from government, commercial, and academic organizations. FIRST aims to foster cooperation and coordination in incident prevention, prompt rapid reaction to incidents, and promote information sharing among members and the community at large. When FIRST was founded in 1990, it had 11 members. By the end of 2001, FIRST consisted of more than 100 response and security teams, which spanned the major global regions.[53]

The Electronic Crimes Task Force (ECTF).[54] The six-year-old ECTF focuses primarily on the New York area, but its network is expanding to include the rest of the United States. The ECTF, a sort of central cyber-crime clearinghouse for all arms of local, state, and national law enforcement, is headed by the New York office of the Secret Service and has a membership of 180 top federal and local law enforcement agencies and prosecutors. The ECTF is careful to guard its top secret data, but it welcomes new members to its network, which consists of about 200 companies from the private sector, mostly from the telecommunications, banking-finance, and vendor-services communities. With the passage of the Patriot Act in 2002, this task force

[51] See Glaessner (1992), Bossone, Promisel (1999).
[52] www.isalliance.org
[53] www.first.org
[54] http://www.ectaskforce.org/

model has been expanded to include the cities of Washington, Boston, Chicago, San Francisco, Miami, and Las Vegas.

InfraGard.[55] InfraGard is a partnership between private industry and the U.S. government, represented by the FBI. The InfraGard initiative was developed to encourage the exchange of information by the government and the private sector. Private sector members and an FBI field representative form local area chapters, which set up their own boards to govern and share information within the membership. Each chapter is also part of the larger InfraGard organization. The NIPC (www.nipc.org), in conjunction with representatives from private industry, the academic community, and the public sector, further developed the InfraGard initiative to expand direct contacts with private sector infrastructure owners and operators and to share information about cyber intrusions, exploited vulnerabilities, and infrastructure threats. The initiative, encouraging the exchange of information by government and private sector members, has continued to expand through the formation of additional InfraGard chapters within the jurisdiction of each FBI field office.

All these arrangements rely on trust, because they make clear that they will not divulge respondents' identities. In some cases, such as with the New York ECTF, partnerships have gone so far as to allow private market participants and law enforcement agencies involved to sign explicit nondisclosure statements as a form of legal safeguard against disclosure of the information being provided. A universally trusted third party collects such information and disseminates it without providing information that could identify the provider, given the possible reputation and other damage related to such a disclosure.

A fruitful exercise might include further study of existing arrangements to share information about electronic security breaches among industry participants, law enforcement, and possibly academic entities with expertise in the technology issues involved. Multilateral lenders such as the World Bank might play a more active role in facilitating such cooperation. In addition, the initiatives of the World Bank and the International Monetary Fund in such areas as initiatives against money laundering and the establishment of financial intelligence units (FIUs) will have to be properly integrated into any well-defined information-sharing framework. For example, suspicious activity reports often can lead to investigations that relate to electronic security breaches and related crimes (e.g., identification thefts).

XI. Pillar VII: Education and Prevention of E-Security Incidents

In many countries throughout the world, statistical analysis reveals that more than 50 percent of electronic security intrusions are carried out by insiders. An uneducated or undereducated workforce is inherently more vulnerable to this type of incident or attack. In contrast, a well-trained workforce, conscious of security issues, can add a layer of protection. Hence, the safety and efficiency of technology is directly related to the training and technical education of the persons using the technology.

That correlation suggests that any effort to reduce and prevent the occurrence of electronic security incidents must rely on an extensive educational effort operating at the following levels: first, the authorities and the persons assigned to examine the financial services

[55] www.nipc.org

providers; second, the systems personnel and others in management at financial services entities; and finally, the users of financial services.

Any plan of action to improve education will need to involve a number of important actions, such as the following:

- Improve awareness and education of financial sector participants about cyber ethics and appropriate user behavior on networked systems. Ensure that employees (and also management), especially those involved in payment system transactions and systems administrators, are aware of the risks and proper approaches to layered security.
- Create institution-wide e-security policies on appropriate behavior and the corresponding channels for reporting intrusions or incidents in close coordination with any effort to improve worldwide information in intrusions (see Section X).
- Develop awareness in the banking community in emerging markets about the need to formulate "incident response plans." In many countries, this will involve efforts to improve capacity; to teach risk assessment, risk management, and prevention; and to develop the essential components of a good security program.
- Facilitate cooperation and transfer of know-how among law enforcement entities, FIUs, and supervisory agencies in developed and emerging markets through such methods as more active exchange programs between personnel. This kind of cooperation can facilitate better education of law enforcement officials, supervisors, and others in emerging market economies about how to deal with e-security.
- Launch some education initiatives in this area targeted to bank examiners, such as at the Toronto Institute, the Federal Reserve courses for bank examiners, or the Financial Stability Institute. The focus of the education should be on techniques for determining whether the layered electronic security systems of brick-and-click banks can be better assessed and evaluated.
- Consider developing a cross-border university outreach program (e.g., involving such entities as Carnegie Mellon's CERT) to promote the training of future e-security professionals, and develop innovative approaches to sharing of information in e-security incidents. Some private entities (e.g., Cisco) provide training at reduced costs for government.
- Develop online programs to improve education of users of e-financial services; develop processes and incentives to have customers report suspicious activities in the use of their accounts. Users and the information they provide are critical to any overall approach to electronic security and risk-sharing.

References

Allen, Julia. 2001. *CERT Guide to System and Network Security Practices.* Indianapolis, Ind.: Addison-Wesley.

Bank of International Settlements. 2001. *Electronic Finance: A New Perspective and Challenges.* BIS Papers No. 7. Basel, Switzerland:

Bator, Francis. 1958. The Anatomy of Market Failure, *QJE.*

Berinato, Scott. 2002. "Finally, a Real Return on Security Spending." *Chief Information Officer (CIO) Magazine.* February 15.

Blake, Seroussi, Smart, *Elliptic Curves in Cryptography*, Cambridge Univ. Press, 1999.

Bossone, Biagio and Larry J. Promisel. 1999. The Role of Self-Regulation in the Financial Sector. The World Bank.

Claessens, Stijn, Thomas Glaessner, and Daniela Klingebiel 2002. *Electronic Finance: A New Approach to Financial Sector Development.* World Bank Discussion Paper No. 431. Washington, D.C.

Claessens, Stijn, Thomas Glaessner, and Daniela Klingebiel. 2001. *E-Finance in Emerging Markets: Is Leapfrogging Possible?* World Bank Financial Sector Discussion Paper No. 7. Washington, D.C.

Cunningham. "Digital Security: Heightened Risks Demand Innovation," *Red Herring*, July 2001.

Federal Trade Commission. 2001. ID Theft. When Bad Things Happen to Your Good Name.

Gilbride, Edward. 2001. *Emerging Bank Technology and the Implications for E-Crime Presentation.* September 3.

Glaessner, Thomas. 1992. External Regulation vs. Self-Regulation: What is the right mix?: An Emerging Markets LAC Perspective. The World Bank.

Internet Security Alliance. Common Sense Guide for Senior Manager. Top Ten Recommended Information Security Practices. 1st Edition. July 2002.

Kahn, Alfred E. 1970. The Economics of Regulation: Principles and Institutions. John Wiley & Sons, Inc.

Konda, Suresh, and Soumyo Moitra. 2000. The Survivability of Network Systems: An Empirical Analysis. Paper. Carnegie Mellon Software Engineering Institute, Pittsburgh, Pa.

OCC Bulletin 2001-35 Examination Procedures to Evaluate Compliance with the Guidelines Safeguarding Customer Information. July 18, 2001.

OCC Alert 2001-4, Network Security Vulnerabilities. April 24, 2001.

OCC Bulletin 2001-12, Bank Provided Account Aggregation Services. February 28, 2001.

OCC Bulletin 2001-8, Guidelines Establishing Standards for Safeguarding Customer Information.

OCC Advisory Letter 2000-12, Risk Management of Outsourced Technology Services. November 28, 2000.

OCC Bulletin 2000-14, Intrusion Risks (May 15, 2000)

OCC Alert 2000-1, Distributed Denial of Service Attacks (February 11, 2000)

United States Department of Justice. 2002. “McNeese” Press Release. Retrieved on March 1, 2002, from http://www.cybercrime.gov/mcneeseArrest.htm.

Schumaker, Troy. 2002. *Cover Your Assets.* Denver, Colorado: North Atlantic Books and Frog Limited.

Shapiro, Carl, and Hal Varian. 1999. *Information Rules: A Strategic Guide to the Network Economy*. Boston, Mass.: Harvard Business School Press.

Soo Hoo, Kevin 2001. "Tangible ROI through Secure Software Engineering." *Secure Business Quarterly*. October.

Sullivan, Bob. 2001. “Massive Credit Heist Fraud Reported.” MSNBC Online. Retrieved on December 22, 2001, from http://www.msnbc.com.

White House. 2000. *Defending America’s Cyberspace: National Plan for Information Systems Protection Version 1.0*. White House, Washington D.C.

Glossary

A -

Abuse of privilege: When a user performs an action that he or she should not have performed according to organizational policy or law.

Access: The ability to enter a secured area, and the process of interacting with a system. Used as either a verb or a noun.

Access authorization: Permission granted to users, programs, or workstations.

Access control: A set of procedures performed by hardware, software, and administrators to monitor access, identify users requesting access, record access attempts, and grant or deny access.

Access-sharing: Permitting two or more users simultaneous access to file servers or devices.

Alphanumeric key: A sequence of letters, numbers, symbols, and blank spaces from one to eighty characters long.

ANSI: The American National Standards Institute. ANSI develops standards for transmission storage, languages, and protocols, and represents the United States in the ISO (International Standards Organization).

Application level gateway [firewall]: A firewall system in which service is provided by processes that maintain complete TCP (telecommunications protocol) connection state and sequencing. Application-level firewalls often readdress traffic so outgoing traffic appears to have originated from the firewall rather than the internal host.

Application logic: The computational aspects of an application, including a list of instructions that tells a software application how to operate.

Audit: The independent collection of records to access their veracity and completeness.

Audit trail: An audit trail may be on paper or on disk. In computer security systems, it is a chronological record of when users log in, how long they are engaged in various activities, what they were doing, and whether any actual or attempted security violations occurred.

Authenticate: In networking, to establish the validity of a user or a communications server.

Authentication: The process of establishing the legitimacy of a node or user before allowing access to requested information. During the process, the user enters a name or account number (identification) and password (authentication).

Authentication tool: A software or hand-held hardware "key" or "token" used during the user authentication process. See *key* and *token*.

Authentication token: A portable device for user authentication. Authentication tokens operate by challenge and response, time-based code sequences, or other techniques that may include paper-based lists of one-time passwords.

Authorization: The process of determining what number of activities is permitted. Usually, authorization is in the context of authentication. Once the user is authenticated, the user may be authorized different levels of access or activity.

Availability: The portion of time a system can be used for productive work, expressed as a percentage.

- B -

Back door: An entry point to a program or a system that is hidden or disguised, often created by the software's author for maintenance. A certain sequence of control characters permits access to the system manager account. If the back door becomes known, unauthorized users (or malicious software) can gain entry and cause damage.

Bandwidth: Capacity of a network or data connection, often measured in kilobits/second (kbps) for digital transmissions.

Bastion host: A system that has been hardened to resist attack at some critical point of entry and that is installed on a network in such a way that it is expected to come under attack. Bastion hosts are often components of firewalls, or may be “outside” Web servers or public access systems. Generally, a bastion host is running some form of general-purpose operating system (LNIX, VMS, WNT, etc.) rather than a ROM-based or firmware operating system.

Biometric access control: Any means of controlling access through human measurements such as fingerprints and iris scans.

Business-critical applications: The vital software needed to run a business, whether custom-written or commercially packaged, such as accounting or finance.

- C -

CERT: The Computer Emergency Response Team, established at Carnegie -Mellon University after the 1988 Internet worm attack named Morris.

Challenge/response: A security procedure in which one communicator requests authentication of another communicator and the latter replies with a preestablished appropriate reply.

Chroot: A technique under UNIX whereby a process is permanently restricted to an isolated subset of the file system.

Client/device: Hardware that retrieves information from a server.

Clustering: A group of independent systems working together as a single system. Clustering technology allows groups of servers to access a single disk array containing applications and data.

Coded file: In encryption, a coded file contains unreadable information.

Combined evaluation: Method using proxy and state or filter evaluations as allowed by administrator. See Stateful evaluation.

Communications server: Procedures designed to ensure that telecommunications messages maintain their integrity and are not accessible by unauthorized individuals.

Computer security: Technological and managerial procedures applied to computer systems to ensure the availability, integrity, and confidentiality of information managed by the computer system.

Computer security audit: An independent evaluation of the controls employed to ensure appropriate protection of an organization's information assets.

Cryptographic checksum: A one-way function applied to a file to produce a unique "fingerprint" of the file for later reference. Checksum systems are a primary means of detecting file-system tampering on UNIX.

- D -

Data-driven attack: A form of attack that is encoded in innocuous-seeming data executed by a user or other software to implement an attack. In the case of firewalls, a data-driven attack is a concern because it may get through the firewall in data form and launch an attack against a system behind the firewall.

Data encryption standard (DES): An encryption standard developed by EBM and then tested and adopted by the National Bureau of Standards. Published in 1977, the DES standard has proven itself over nearly 20 years of use in both government and private sectors.

Decode: Conversion of encoded text to plain text through the use of a code.

Decrypt: Conversion of either encoded or enciphered text into plain text.

Dedicated: A special-purpose device. Although capable of performing other duties, it is assigned to only one.

Defense in depth: The security approach whereby each system on the network is secured to the greatest possible degree. May be used in conjunction with firewalls.

DES: Data encryption standard.

DNS spoofing: Assuming the Domain Name Server (DNS) name of another system by either corrupting the name service cache of a victim system or compromising a domain name server for a valid domain.

Dual-homed gateway: (1) A system that has two or more network interfaces, each of which is connected to a different network. In firewall configurations, a dual-homed gateway usually acts to block or filter some or all of the traffic trying to pass between the networks. (2) A firewall implement that does not use a screening router.

- E -

E-mail bombs: Code that when executed sends many messages to the same address for the purpose of using up disk space or overloading the e-mail or Web server.

Encrypting router: See Tunneling router and Virtual network perimeter.

Encryption: The process of scrambling files or programs, changing one character string to another through an algorithm (such as the DES algorithm).

End-to-end encryption: Encryption at the point of origin in a network, followed by decryption at the destination.

Environment: The aggregate of external circumstances, conditions, and events that affect the development, operation, and maintenance of a system.

ERP (enterprise resource planning): ERP systems permit organizations to manage resources across the enterprise and completely integrate manufacturing systems.

Extranet: Extranet refers to extending the LAN via remote or Internet access to partners outside your organization, such as frequent suppliers and purchasers. Such relationships should be over an authenticated link to authorized segments of the LAN and are frequently encrypted for privacy.

- F -

Fat client: A computing device, such as a PC or Macintosh, that includes an operating system, RAM, ROM, a powerful processor, and a wide range of installed applications that can execute on the desktop or 100 percent on the server under a server-based computing architecture. Fat clients can operate in a server-based computing environment.

Fault tolerance: A design method that ensures continued systems operation in the event of individual failures by providing redundant system elements.

Firewall: A system or combination of systems that enforces a boundary between two or more networks.

Flooding programs: Implementing a code that when executed will bombard the selected system with requests in an effort to slow down or shut down the system.

Anonymous FTP [Define acronym]: A guest account that allows anyone to login to the FTP server. It can be a point to begin access on the host server.

- G -

Gateway: A bridge between two networks.

Generic utilities: General purpose code and devices—that is, screen grabbers and sniffers that look at data and capture such information as passwords, keys, and secrets.

Global security: The ability of an access-control package to permit protection across a variety of mainframe environments, providing users with a common security interface to all.

GPS (global positioning system) : Used primarily for navigation, this satellite-based system maps the location of various receivers on earth.

Granularity: The relative fineness or coarseness by which a mechanism can be adjusted.

GSM: Groupe Spécial Mobile, the European Union's digital cellular standard.

- H -

Hack: Any software in which a significant portion of the code was originally another program.

Hackers: Those intent on entering an environment to which they are not entitled entry for whatever purpose (e.g., entertainment, profit, theft, prank), usually involving iterative techniques, escalating to more advanced methodologies, and use of devices to intercept the communications property of another.

Host-based security: The technique of securing an individual system from attack. Host-based security is operating system- and version-dependent.

Hot standby: A backup system configured in such a way that it may be used if the system goes down.

Hybrid gateway: An unusual configuration with routers that maintain the complete state of the TCP/IP connections or examine the traffic to try to detect and prevent attack (may involve host). If very complicated, it is difficult to attach, maintain, and audit.

- I -

ICA: An acronym for Citrix's Independent Computing Architecture, a three-part server-based computing technology that separates an application's logic from its user interface and allows 100 percent application execution on the server.

IETF (The Internet Engineering Task Force): A public forum that develops standards and resolves operational issues for the Internet. IETF is purely voluntary.

Information systems technology: The protection of information assets from accidental or intentional but unauthorized disclosure, modification, or destruction or the inability to process that information.

Insider attack: An attack originating from inside a protected network.

Internet: A web of different, intercommunicating networks funded by both commercial and government organizations. The Internet had its roots in early 1969 when the ARPANET was formed. ARPA stands for Advanced Research Projects Agency (which was part of the U.S. Department of Defense). One of the goals of ARPANET was research in distributed computer systems for military purposes. The first configuration involved four computers and was designed to demonstrate the feasibility of building networks using computers dispersed over a wide area. The advent of open networks in the late 1980s required a new model of communications. The amalgamation of many types of systems into mixed environments demanded a better translator between these operating systems and a nonproprietary approach to networking in general. Telecommunications Protocol/Internet Protocol (TCP/IP) provided the best solutions.

Intrusion detection system: A system dedicated to the detection of break-ins or break-in attempts manually either via software expert systems that operate on logs or other information available on the network.

IP sniffing: Stealing network addresses by reading the packets. Harmful data is then sent stamped with internal trusted addresses.

IP splicing: An attack whereby an active, established session is intercepted and co-opted by the attacker. EP splicing attacks may occur after an authentication has been made, permitting the attacker to assume the role of an already authorized user. Primary protections against IP splicing rely on encryption at the session or network layer.

IP spoofing: An attack whereby a system attempts to illicitly impersonate another system by using its EP network address.

ISO (International Standards Organization): Sets standards for data communications.

ISSA: Information Systems Security Association.

- J -

- K -

Key: In encryption, a sequence of characters used to encode and decode a file. One can enter a key in two formats: alphanumeric and condensed (hexadecimal). In the network access security market, "key" often refers to the "token," or authentication tool, which is a device used to send and receive challenges and responses during the user authentication process. Keys may be small, hand-held hardware devices similar to pocket calculators or credit cards or they may be loaded onto a PC as copy-protected software.

- L -

Least privilege: Designing operational aspects of a system to operate with a minimum amount of system privilege. This design reduces the authorization level at which various actions are performed and decreases the chance that a process or user with high privileges may be caused to perform unauthorized activity resulting in a security breach.

Local area network (LAN): An interconnected system of computers and peripherals; LAN users share data stored on hard disks and can share printers connected to the network.

Logging: The process of storing information about events that occurred on the firewall or network.

Log processing: How audit logs are processed, searched for key events, or summarized.

Log retention: How long audit logs are retained and maintained.

- M -

Mobile code: A program downloaded from the Internet that runs automatically on a computer with little or no user interaction.

Multi-user capability: The ability for multiple concurrent users to log on and run applications from a single server.

- N -

Network computer (NC): A "thin" client hardware device that executes applications locally by downloading them from the network. NCs adhere to a specification jointly developed by Sun, IBM, Oracle, Apple, and Netscape. NCs typically run Java applets within a Java browser or Java applications within the Java Virtual Machine.

Network computing architecture: A computing architecture in which components are dynamically downloaded from the network into the client device for execution by the client. The Java programming language is at the core of network computing.

Network-level firewall: A firewall in which traffic is examined at the network protocol packet level.

Network worm: A program or command file that uses a computer network as a means for adversely affecting a system's integrity, reliability, or availability. A network worm may attack from one system to another by establishing a network connection. The worm is usually a selfcontained program that does not need to attach itself to a host file to infiltrate network after network.

NIPC (National Infrastructure Protection Center): NIPC brings together representatives from U.S. government agencies, state and local governments, and the private sector in a partnership to protect the nation's critical infrastructures. NIPC's mission is to serve as the U.S. government's focal point for threat assessment, warning, investigation, and response in cases of threats or attacks against electronic critical infrastructures.

- O -

One-time password: In network security, a password issued only once as a result of a challenge-response authentication process. Cannot be "stolen" or reused for unauthorized access.

Operating system: System software that controls a computer and its peripherals. Modern operating systems, such as Unix, Linux, and Windows XP handle many of a computer's basic functions.

Orange book: The Department of Defense Trusted Computer System Evaluation Criteria. It provides information to classify computer systems, defining the degree of trust that may be placed in them.

- P -

Password: A secret code assigned to a user, known by the computer system. Knowledge of the password associated with the user ID is considered proof of authorization. (See One-time password.)

Performance: A major factor in determining the overall productivity of a system, performance is primarily tied to availability, throughput, and response time.

Perimeter-based security: The technique of securing a network by controlling access to all entry and exit points of the network.

PIN (personal identification number): In computer security, a PIN is known only to the user and used during the authentication process. (See Challenge/response; Two-factor authentication.)

Policy: Organizational-level rules governing acceptable use of computing resources, security practices, and operational procedures.

Private key: The element of a public/private key pair that is kept secret by the key pair owner. The private key is used to decrypt messages that have been encrypted by the corresponding public key. It also is used to construct a digital signature – the document to be signed first is hashed using a secure hash algorithm; then encrypting the hashed value using the private key forms the digital signature.

Protocols: Agreed-on methods of communications used by computers.

Proxy: (1) A method of replacing the code for service applications with an improved version that is more security-aware. Preferred method is by "service communities" rather than individual applications. Evolved from socket implementations. (2) A software agent that acts on behalf of a user. Typical proxies accept a connection from a user, make a decision as to whether the user or client IP address is permitted to use the proxy, perhaps does additional authentication, and then completes a connection on behalf of the user to a remote destination.

Public key: - The element of a public/private key pair that can be known by anyone. The public key is used to encrypt information that is to be intelligible only to the holder of the corresponding private key. It also is used to decrypt a digital signature in order to compare the decrypted digital signature and the hashed value of the signed document.

Q -

- R -

Remote access: The hookup of a remote computing device via communications lines, such as ordinary phone lines or wide area networks, to access network applications and information.

Remote presentation services protocol: A protocol is a set of rules and procedures for exchanging data between computers on a network. A remote presentation services protocol transfers user interface, keystrokes, and mouse movements between a server and a client.

Risk analysis: The analysis of an organization's information resources, existing controls, and computer system vulnerabilities. It establishes a potential level of damage in dollars or other assets.

Rogue program: Any program intended to damage programs or data. Encompasses malicious Trojan horses.

RSA: A public key cryptosystem named by its inventors—Rivest, Shamir, and Adelman—who hold the patent.

- S -

Salami slice: A hacker method for the acquisition of funds. A database of account information is copied. Then on a later date all accounts are charged a minimal amount, so as not to arouse suspicion.

Scalability: The ability to expand a computing solution to support large numbers of users without having an impact on performance.

Screened host gateway: A host on a network behind a screening router. The degree to which a screened host may be accessed depends on the screening rules in the router.

Screened subnet: An isolated subnet created behind a screening router to protect the private network. The degree to which the subnet may be accessed depends on the screening rules in the router.

Screening router: A router configured to permit or deny traffic using filtering techniques; based on a set of permission rules installed by the administrator. A component of many firewalls usually used to block traffic between the network and specific hosts on an IP port level. Not very secure; used when speed is the only decision criterion.

Symmetric key: The secret key used for both encryption and decryption with a symmetric cipher such as DES, triple DES, or AES.

Server: The control computer on a local area network that controls software access to workstations, printers, and other parts of the network.

Server-based computing: An innovative, server-based approach to delivering business-critical applications to end-user devices, whereby an application's logic executes on the server and only the user interface is transmitted across a network to the client. Its benefits include single-point management, universal application access, bandwidth-independent performance, and improved security for business applications.

Server farm: A group of servers that are linked together as a "single system image" to provide centralized administration and horizontal scalability.

Session shadowing: A feature of Citrix WinFrame and MetaFrame that allows administrators and technical support staff to join remotely or take control of a user's session for diagnosis, support, and training.

Session stealing: See IP splicing.

Single-point control: Helps to reduce the total cost of application ownership by enabling applications and data to be deployed, managed, and supported at the server. Single-point control enables application installations, updates, and additions to be made once, on the server, and then instantly made available to users anywhere.

Smart card: A credit card–sized device with embedded microelectronics circuitry for storing information about an individual. This is not a key or token, as used in the remote access authentication process.

Social engineering: An attack based on deceiving users or administrators at the target site. Social engineering attacks are typically carried out by telephoning users or operators and pretending to be an authorized user to attempt to gain illicit access to systems.

Stateful evaluation: Methodology using mixture of proxy or filtering technology intermittently, depending on perceived threat (or need for speed.)

- T -

TCO (total cost of ownership): A model that helps IT professionals understand and manage the budgeted (direct) and unbudgeted (indirect) costs incurred for acquiring, maintaining, and using an application or a computing system. TCO normally includes training, upgrades, and administration as well as the purchase price. Lowering TCO through single-point control is a key benefit of server-based computing.

Thin client: A low-cost computing device that works in a server-centric computing model. Thin clients typically do not require state-of-the-art, powerful processors and large amounts of RAM and ROM because they access applications from a central server or network. Thin clients can operate in a server-based computing environment.

Token: In authentication, a device used to send and receive challenges and responses during the user authentication process. Tokens may be small, hand-held hardware devices similar to pocket calculators or credit cards. See Key.

Trojan horse: (1) Any program designed to do things the user of the program did not intend to do or that disguise its harmful intent. (2) A program that installs itself while the user is making an authorized entry, and then is used to break in and exploit the system.

Tunneling router: A router or system capable of routing traffic by encrypting it and encapsulating it for transmission across an untrusted network for eventual de-encapsulation and decryption.

Turn commands: Commands inserted to forward mail to another address for interception.

Two-factor authentication: Two-factor authentication is based on something a user knows (factor one) plus something the user has (factor two). In order to access a network, the user must have both "factors," just as he or she must have an ATM card and a PIN to retrieve money from a bank account. In order to be authenticated during the challenge and response process, users must have this specific (private) information.

- U -

User: Any person who interacts directly with a computer system.

User ID: A unique character string that identifies a user.

User identification: User identification is the process by which a user identifies herself to the system as a valid user—as opposed to authentication, which is the process of establishing that the user is indeed that user and has a right to use the system.

User interface: The part of an application that the user works with. User interfaces can be text-driven, such as DOS, or graphical, such as Windows.

- V -

VPN (virtual private network): A private connection between two machines that sends private data traffic over a shared or public network, such as the Internet. VPN technology lets an organization securely extend its network services over the Internet to remote users, branch offices, and partner companies.

Virtual network perimeter: A network that appears to be a single protected network behind firewalls, but actually encompasses encrypted virtual links over untrusted networks.

Virus: A self-replicating code segment. Viruses may or may not contain attack programs or trapdoors.

- W -

WEP (Wireless Equivalent Protocol): A protocol designed to be implemented over WLANs to offer the same security features as a physical wire: confidentiality, access control, and data integrity.

Windows-based terminal (WBT): A fixed-function thin-client device that connects to a Citrix WinFrame or MetaFrame server and terminal server to provide application access. The key differentiator of a WBT from other thin devices is that all application execution occurs on the server; there is no downloading or local processing of applications at the client.

WLAN (wireless local area network): A wireless Network that corresponds to wireless laptops.

- XYZ -

Y2K: An acronym for the year 2000 problem, which involved three issues: two-digit data storage, leap-year calculations, and special meanings for dates.

Annex I. Risk Management: A Blueprint for Layered Security[56]

Introduction

As financial entities increase their reliance on open architecture network systems, the financial impact from computer-related attacks and the subsequent disruption of services provided by core digital infrastructures will increase exponentially. In today's business-model climate, the gap between the risk management of physical assets and the risk management of informational assets is enormous. Moreover, the gap between a bank's operational risk management and information technology risk management requirements places most if not all of the bank's assets at risk. This annex sets forth an electronic blueprint that illustrates the methodology, procedures, and technologies needed to "layer security" around the e-financial online platforms of today.

Section A of this annex discusses the paradigm of *risk analysis*. Section B highlights the appropriate best practices for access controls. Section C identifies and critiques the various types of firewalls and how they should be configured. Section D reviews the greatest content- and protocol-related threats to network systems. Section E highlights the effectiveness of the three classes of intrusion detection systems. Section F illustrates the best practices to prevent outbreaks of viruses within networks. Section G covers encryption. Section H identifies the administrative safeguards that best secure modern servers. Section I discusses the need for vigilance through proactive penetration testing of networks so as to discern a vulnerability before the criminals do. Section J highlights the proper procedures for secure systems administration. Section K sets guidelines for a corporate-wide reaction plan in case the network is damaged or destroyed following a hack attack.

A. Risk Management

Good business and good regulation require that financial institutions manage risk appropriately. Security is a twofold process. The first part is risk analysis, which has three major components: identify and inventory assets for a baseline, analyze and assign values to the assets, and establish how critical each asset is, in priority order. The second part of security is development of an approach to risk management. The major elements of risk management are to develop and implement policies and procedures, educate users (employees and customers), and audit and monitor for quality assurance. A prudent approach might reflect the following thesis: "Expect to be hit – Prepare to survive."

[56] The annexes to this paper were created to be a comprehensive blueprint for adequate, layered, electronic security. The contributors and reviewers came from the e-security industry, the law enforcement community, and academic institutions. The reviewers of these annexes believe that the procedural guidelines for e-security outlined in the annexes are state-of-the-art as of May 2002. These reviewers were Dave Thomas, Chief Investigator for the FBI Computer Crimes Division; Keith Schwalm, Director of Infrastructure Protection, President's Critical Infrastructure Protection Board, Rick Fleming, Vice President of Security Operations for Digital Defense Inc.; Linda McCarthy, Vice President of Systems Engineering at Recourse Technologies Inc.; Simon Martinez, Director of the Federal Deposit Insurance Corporation's Computer Security Incident Response Team; Gary Sullivan, President of Galaxy Computer Services Inc.; Chris Bateman, Technical Analyst for the Computer Emergency Response Team; Dorothy Denning, Professor of Computer Engineering at Georgetown University; Joseph Pelton, Executive Director of the Clarke Institute; Stephanie Lanz, Visiting Research Scholar at the Center for Emerging Threats and Opportunities at Quantico VA.; John Frazzini, Special Agent for the Secret Service Financial Crimes Division, and Bill Worley, HP fellow, Chief Scientist for Hewlett Packard, Inc.

The three general axioms to remember in building a security program are as follows:

- Attacks and losses are inevitable.
- Security buys time.
- The network is only as secure as its weakest link.

Effective risk management of electronic activity has three phases, as follows.[57]

Phase 1:[58] Build asset-based threat profiles. An analysis team consisting of management and technical staff must identify and inventory the institution's mission-critical assets, set priorities among them, and determine what is being done to protect these assets. The creation of an Information Security Officer is critical to the management of risk within banks network. It is essential that this position be filled so as to bring accountability into the process of electronic security. The Information Security Officer should be responsible for carrying out each layer of the 11-step risk management framework. Trained in OCTAVE he or she will be the focal point for proactive e-risk mitigation.

Phase II: Identify infrastructure vulnerabilities and interdependencies. A multidisciplinary team analyzes key system vulnerabilities, exploits interdependencies, and identifies potential points of entry and workflow processes. An outside contractor then should initiate a formal assessment and design a secure architecture. The architecture should include, at a minimum, appropriate policies, procedures, protocols, and tools to evaluate the current software and hardware and to audit and monitor against the workflow processes and procedures.

Phase III: Develop a business continuity plan. A multidisciplinary team should identify potential attack scenarios and develop contingency plans to eliminate/mitigate risk if any of the mission-critical assets identified in Phase II are compromised. Appropriate policies and procedures should be in place company-wide, and periodic testing and surprise attacks should be mounted to assure management that safeguards are in place, up-to-date, and in use.

The Ghost Site

As part of their business continuity process, Banks should develop an off-line site as a preemptive measure to an interruption. This site should be fully functional and have the ability to "go live within minutes. The robustness of the ghost site is directly related to the business needs of the organization. At a minimum, the ghost site should provide appropriate communications to customers and be supported by a call center. This ghost site should consider all aspects of the business: Customers vendors, public relations, legal, marketing, investor relations etc. It is imperative that its development team is designed from a cross-section of the organization. Discontinuity will happen to an organization and this exercise ensures management that they have a measured response for mitigating risk.

The BITS Financial Services Security Laboratory is a testing facility. It was launched in July 1999 to test products and services that strengthen the security of electronic payment systems and related e-commerce technologies. The criteria established by BITS need to be strengthened. The annexes to this paper serve to depict the minimum technological baseline that banks should incorporate into their information technology (IT) security. Based on both the BITS

[57] Operationally Critical Threat, Asset, and Vulnerability Evaluation, and its acronym OCTAVE, are service marks of Carnegie Mellon University. www.cert.org

[58] OCTAVE Method Implementation Guide v2.0, Computer Emergency Response Team, Carnegie Mellon University.

recommendations (see box 1 in this annex) and the 14 recommendations from the Electronic Banking Group (see box 5 in Section VII of the paper), these annexes provide a blueprint for "true" layered security.

Box 1. BITS Master Security Criteria

- **Identification:** The system shall have the capability of associating a user with an unambiguous identifier by which the said user shall be held accountable for the actions and events initiated by that user.
- **Nonrepudiation:** The system shall have the capability of preventing users from successfully denying actions and events of users acting in the role of a sender or receiver.
- **Authentication:** The system shall offer features to verify the claimed identity of a user before allowing system access to the said user.
- **Authorization:** The system shall offer features to support the following restrictions:
  - No user shall be allowed access to the system without identification and authentication;
  - No user shall be allowed access to a resource of the system unless specifically authorized.
- **Confidentiality:** The system shall offer features to ensure that sensitive information shall be communicated and stored in a way such that only authorized users are allowed access.
- **Data Integrity:** The system shall offer features to ensure that either:
  - The data shall not be modified or altered without authorization in either storage or in transit; or
  - Any unauthorized modification of data shall yield an auditable security-related event.
- **Audit:** The system shall offer features to support the following functions:
  - Maintain a history file (e.g., audit log) that records all security-related events pertinent to establishing an audit trail for a “post-mortem” analysis of a suspected security breach;
  - Ensure integrity of the audit log;
  - Generate customized audit reports;
  - Protect audit log from unauthorized[un okay?] access;
  - Support administrator-selectable alerts for specified security related events;
  - Support audit records of administrative events.
- **Data Disposal:** The system shall ensure that there is no residual data exposed to unauthorized users as resources are allocated to those data objects or released from those data objects.
- **System Integrity:** The system shall offer features to support the following functions:
  - Perform integrity checks for system function;
  - Retain the security parameters after the occurrence of events such as system restart, disaster recovery, arrival of sensitive dates, et cetera;
  - Provide the backup capability to restore the system, when necessary, to a well-defined state;
  - Ensure the security features are always invoked and may not be bypassed unless authorized and configured to do so.
- **Security Administration:** The system shall offer features to selectively authorize a highly privileged user (as security administrator) to perform day-to-day activities.
- **Guidance:** The vendor shall supply the following product support capability:
  - A cogent security-related document for administration that would be made available as a hard copy or an electronic file, as an entity unto itself, and not fragmented throughout the manuals.

B. Access Controls—Authentication

The first line of defense and the most inexpensive technique is the use of access controls and the implementation of policies for computer usage. At a minimum, any financial institution that chooses to implement passwords as the primary access control mechanism to the network should mandate the following best practices. (See Annex II for more detailed analysis of access controls.)

1. Users should be required to issue both a user ID and a password at the time of logon.
2. File and directory access control should be set according to the sensitivity and use of the files and directories.
3. Users should be granted rights and privileges to available system resources only on a need-to-know, need-to-use basis.
4. The host system should be able to identify both the workstation and the workstation connection point (i.e., location) at logon.
5. Data access control should be determined at the appropriate level in each division.
6. File access privileges should be identified as read, read-only, write (with separate add and update levels), execute, execute-only, create, rename, delete, change access, and none.
7. System resource access should be assignable on an individual, group, or public basis.
8. The user ID format should be seven (7) characters. The remaining characters should be alphanumeric special characters (e.g., *, %, @), which are even more secure than numeric. This format should be applied to all platforms the user is authorized to access.
9. System passwords should be a minimum of 6 up to a maximum of 16 characters long.
10. System passwords should be changed at least every 90 calendar days.
11. System passwords should not be reused within the last 10 passwords.
12. There should be no echoing of password positions, actual characters used, or dummy characters (e.g., display of asterisk for each password character position).
13. New system passwords should always be entered twice for verification.
14. Attempts by general users to enter a correct password should be limited to five (5) attempts, after which the system will disable the account.

C. Firewalls

There are three types of firewalls: packet filter, stateful inspection, and application proxy. A packet filter firewall works by looking at each packet of network information and determining, based on the contents of the headers of that packet, whether it is allowed to traverse the network. A stateful inspection firewall also looks at the packets, but instead of looking at just the addressing information, it looks at how the connection has been set up between the computers to determine if the packet of information is in a valid state. For instance, after the connection is set up in a stateful inspection firewall and a packet of network traffic shows up destined to a particular host, the firewall can look to see if that connection has been established. If not, the packet may contain spoofed information and it should be discarded. An application proxy firewall is a process-based control device that stands guard in front of the application to permit only authorized parties access to all or parts of an application.

Firewalls have one basic goal: to keep network traffic from passing a given point that does not meet certain connection criteria. For instance, it is very common for firewalls to only allow traffic into servers connected on ports 80 (http) and 443 (https). While this may stop much unwanted traffic, additional steps are needed to protect the resources. It is very important to remember that the firewall *does not* stop traffic on the ports that *are* allowed. A firewall cannot prevent what is already allowed through the system.

Recommendations for Proper Firewall Configuration

1. The firewall should be placed in between your network router and your network or given application.

2. It is important to minimize and limit any network protocol that is not required by your organization. As an example, if you are not using Novell, then IPX should not be allowed to traverse your network.
3. Routers need proper configuration. The most common mistake in configuring firewalls is to allow servers that should accept only inbound connections—for example, Web servers—to make outbound connections. This is important because if your Web server is compromised as a result of a security breach and the firewall is configured to not allow that server to initiate a connection to another system, then the hacker has found a dead-end connection. Effectively, the hacker cannot further compromise your systems using that Web server.[59] Routers are capable of filtering packets as well. According to Peter Tippett, Chief Technical Officer of Trusecure Inc.: "Over 92 percent of routers have turned their default feature off. This feature, which exists in all Cisco routers, is typically disallowed due to the existence of an Any/Any rule that turns this safety feature off. If the rule for the router is changed to Default/Deny, your router becomes, in essence, an extra packet firewall, thus reducing risk to your business by 2,000 percent."
4. Firewalls need to remain current. They should be updated like virus scanners and patches at regular intervals, when a serious vulnerability has been discovered, and when a patch is available.
5. Ingress and egress filtering should be used. Spoofing Internet protocol (IP) addresses is a common method used by attackers to hide their tracks when they attack a victim. For example, the very popular smurf attack uses a feature of routers to send a stream of packets to thousands of machines. Each packet contains a spoofed source address of a victim. The computers to which the spoofed packets are sent flood the victim's computer, often shutting down the computer or the network. Performing filtering on traffic coming into your network (ingress filtering) and going out (egress filtering) can help provide a high level of protection. The filtering rules60 are as follows:
   - Any packet coming into your network must not have a source address of your internal network.
   - Any packet coming into your network must have a destination address of your internal network.
   - Any packet leaving your network must have a source address of your internal network.
   - Any packet leaving your network must not have a destination address of your internal network.
   - Any packet coming into your network or leaving your network must not have a source or destination address of a private address or an address listed in RFC1918 reserved space. These include 10.x.x.x/8, 172.16.x.x/12, or 192.168.x.x/16, and the loopback network 127.0.0.0/8.
   - Block any source-routed packets or any packets with the IP options field set.
   - Reserved, DHCP auto-configuration, and Multicast addresses should also be blocked:
     - 0.0.0.0/8
     - 169.254.0.0/16
     - 192.0.2.0/24
     - 224.0.0.0/4
     - 240.0.0.0/4
6. Rate-limiting filters should be employed.

[59] Contributed by Rick Fleming, Vice President of Security Operations for Digital Defense Inc.
[60] Input provided by the SANS Institute.

7. If users are allowed to connect from the Internet to the internal network, the access should be restricted to either a virtual private network (VPN) or an encrypted software session.[61]
8. Security policies should be put in place that control both internal and external access to the network. Security policies should not allow the needs of IT and specific applications to dictate—for example, if one needs to access a system internally using a wireless laptop and the systems administrator places the access point outside of the firewall. This is a configuration error. As recent Nimda and Code Red worms taught us, a firewall reduces the number of hack attempts, but it does not eliminate them.
9. The list of what a systems administrator should allow/disallow through the firewall depends on many factors. The critical factor is where on the network the firewall sits. For example, a firewall positioned between a Web server and an application layer will pass one set of protocols, and a DMZ firewall that sits in front of the mail server will pass a completely different set. In either case, the system should be designed to allow through a limited set of well-behaved telecommunications protocol (TCP) ports.[62] However, if the firewall is guarding the corporate network, only authenticated protocols should be given access to bolster security. And, when configuring a firewall, any protocol that is not required for the organization to function properly should be disallowed. Moreover, proper configuration dictates that only the minimum set of protocols be allowed access to the minimum set of specific hosts that are needed to run the network. The system denies all and logs those packets that are dropped. By blocking traffic to these ports at the firewall or other network perimeter protection device, you add an extra layer of defense that helps protect you from configuration mistakes. Note, however, that using a firewall to block network traffic directed to a port does not protect the port from disgruntled coworkers who are already inside your perimeter or from hackers who may have penetrated your perimeter using other means. According to the SANS Institute, the following ports should be blocked:[63]
   - Login services—telnet (23/tcp), SSH (22/tcp), FTP (21/tcp), NetBIOS (139/tcp), rlogin et al. (512/tcp through 514/tcp)
   - RPC and NFS—Portmap/rpcbind (111/tcp and 111/udp), NFS (2049/tcp and 2049/udp), lockd (4045/tcp and 4045/udp)
   - NetBIOS in Windows NT—135 (tcp and udp), 137 (udp), 138 (udp), 139 (tcp). Windows 2000—earlier ports plus 445(tcp and udp)
   - X Windows—6000/tcp through 6255/tcp
   - Naming services—DNS (53/udp) to all machines that are not DNS servers, DNS zone transfers (53/tcp) except from external secondaries, LDAP (389/tcp and 389/udp)
   - Mail—SMTP (25/tcp) to all machines that are not external mail relays, POP (109/tcp and 110/tcp), IMAP (143/tcp)
   - Web—HTTP (80/tcp) and SSL (443/tcp) except to external Web servers; may also want to block common high-order HTTP port choices (8000/tcp, 8080/tcp, 8888/tcp, etc.)
   - "Small Services"—ports below 20/tcp and 20/udp, time (37/tcp and 37/udp)
   - Miscellaneous—TFTP (69/udp), finger (79/tcp), NNTP (119/tcp), NTP (123/udp), LPD (515/tcp), syslog (514/udp), SNMP (161/tcp and 161/udp, 162/tcp and 162/udp), BGP (179/tcp), SOCKS (1080/tcp)

[61] Ibid.
[62] Contributed by Peter Penfield of Cisco Systems.
[63] Refer to www.sans.org/top20.htm.

- ICMP—–block incoming echo request (ping and Windows traceroute); block outgoing echo replies.

In addition, when using a firewall, an organization should consider the type of technology used and determine whether that technology is sufficient, given the new demands and uses on it. For example, an organization that is simply accessing the Internet, but is not providing any Internet-accessible services, might be adequately protected through the use of network address translation and filtering. Another organization might require a firewall that supports application proxies for additional protection.[64] In many cases the best firewall to select is one that combines multiple forms of firewall technology. For example, both the Cisco PIX and Watchguard Firebox systems effectively combine stateful inspection with application level proxies for protection.

The cost-effectiveness of a firewall for an institution depends on four factors: the type of connectivity, the amount of bandwidth, the number of gateways, and the skills set of the firewall and systems administrator. The firewall should be able to support the type of connection the institution has, for example, and the staff needs to be able to administer it effectively. Generally, the best option is a good hardware-based firewall that is dedicated to firewall functionality and is a true “stateful” firewall. Which firewall the institution uses depends greatly on the institution's needs. Implementing a $10,000 firewall such as Secure Computing’s Sidewinder or Cisco’s PIX into an organization with only 50 employees will not provide a good return on investment. At the same time, a smaller class of firewall would be insufficient for an organization with more than 1,000 employees.

D. Active Content Filtering

Once connected to the Internet, an individual or an organization undertakes a degree of risk from malicious code. The impact of allowing malicious content to enter a banking network unchallenged is far-reaching. A well-defined policy and procedure detailing proper network usage should be implemented institution-wide, and implementation should include active content filtering. The following section depicts what active content necessitates controlled access.

Filter all executable attachments; attachments that could affect the computer’s registry and many others attached to emails should be filtered out.

At the browser level, it is necessary and prudent to filter all material that is defined by policy to be inappropriate for the workplace or that is contrary to the institution's established workplace policies. Active content filtering can be accomplished by filtering Internet sites that would otherwise be accessible to employees and through the use of e-mail filters that are designed to detect references to a company’s proprietary data.[65] All content that does not promote the functioning of an organization should be filtered. Technology can be deployed to detect references to the company's proprietary data in e-mail messages and their attachments.

A list of the greatest content- and protocol-related threats to network systems follows.

1. Hostile Active X
2. Javascript
3. Remote Procedure Calls (RPCs)
4. Perimeter-based security (PBS)

[64] Contributed by Galaxy Computer Services Inc.

[65] Input provided by Galaxy Computer Services Inc.

5. Berkeley Internet Name Domain (BIND): The following steps should be taken to defend against the BIND vulnerabilities:[66]
   - Disable the BIND name daemon (called "named") on all systems that are not authorized to be DNS servers. Some experts recommend you also remove the DNS software.
   - On machines that are authorized DNS servers, update to the latest version and patch level. Use the guidance contained in the following advisories:
     For the NXT vulnerability: http://www.cert.org/advisories/CA-99-14-bind.html. For the QINV (Inverse Query) and NAMED vulnerabilities: http://www.cert.org/advisories/CA-98.05.bind_problems.html http://www.cert.org/summaries/CS-98.04.html.
   - Run BIND as a nonprivileged user for protection in the event of future remote-compromise attacks. (However, only processes running as root can be configured to use ports below 1024—a requirement for DNS. Therefore you must configure BIND to change the user ID after binding to the port.)
   - Run BIND in a root directory structure for protection in the event of future remote-compromise attacks.
   - Disable zone transfers except from authorized hosts.
   - Disable recursion and glue fetching to defend against DNS cache poisoning.
   - Hide your version string.
6. Simple Network Management Protocol (SNMP) default community strings[67]:
   - If you do not absolutely require SNMP, disable it.
   - If you must use SNMP, use the same policy for community names as used for passwords. Make sure that they are difficult to guess or crack, and that they are changed periodically.
   - Validate and check community names using snmpwalk. Additional information can be found at http://www.zend.com/manual/function.snmpwalk.php.
   - Filter SNMP (Port 161/UDP) at the border-router or firewall unless it is absolutely necessary to poll or manage devices from outside of the local network.
   - Where possible, make MIBs read-only.
7. JVM vulnerability[68]
8. Sendmail[69]—The following steps should be taken to protect Sendmail: Upgrade to the latest version of Sendmail and/or implement patches for Sendmail; http://www.cert.org/advisories/CA-97.05.sendmail.html. Do not run Sendmail in daemon mode (turn off the -bd switch) on machines that are neither mail servers nor mail relays.
9. Internet Message Access Protocol (IMAP) and Post Office Protocol (POP)
10. Sadmind and mountd: Sadmind allows remote administration access to Solaris systems, providing a graphical user interface for system administration functions. Mountd controls and arbitrates access to NFS mounts on UNIX hosts. The following actions[70] will protect against NFS vulnerabilities, including sadmind and mountd:
    - Wherever possible, turn off and/or remove sadmind and mountd on machines directly accessible from the Internet.
    - Install the latest patches:

[66] The SANS Institute, www.sans.org/top20.htm.

[67] For more comprehensive information regarding this vulnerability refer to: http://www.cert.org/advisories/CA-2002-03.html

[68] There is a JVM patch that will prevent the copying of Web sites visited as well as the copying of passwords typed.

[69] Sendmail is the program that sends, receives, and forwards most electronic mail processed on UNIX and Linux computers.

[70] The SANS Institute, www.sans.org/top20.htm.

For Solaris Software, http://sunsolve.sun.com;
For IBM AIX Software,
http://techsupport.services.ibm.com/support/rs6000.support/downloads,
http://techsupport.services.ibm.com/rs6k/fixes.html;
For SGI Software, http://support.sgi.com/;
For Compaq (Digital UNIX), http://www.compaq.com/support.

- Use host/IP-based export lists.
- Set up export file systems for read-only or no suid where possible.
- Use nfsbug to scan for vulnerabilities.

11. E-mail
    - Filter all .exe attachments.
    - Filter all .doc attachments.
    - Consider filtering all arriving and departing e-mail by a spam threshold (greater than 40 identical messages blocked and source traced, if inside the network).

E. Intrusion Detection Systems

"An intrusion should be defined as any system/network activity that cannot be justifiably explained or activity that results in the disruption of services or loss of data."[71]

An intrusion is a suspicious pattern that may indicate a network or system attack from someone attempting to break into or compromise a system.

Intrusion detection systems (IDSs) need to be monitored 24 hours a day to obtain the best return on investment. This work schedule positions the institution to respond to suspected intrusions in a timely manner and prevents the inadvertent loss of resources resulting from a misconfigured IDS.[72]

In addition, an IDS should be installed inside the firewall and should operate only on traffic that is not allowed within the firewall's perimeter. An IDS installed inside the firewall is critical for assisting the administrator to determine if the attack was successful in breaching the firewall.

What type of monitoring approach is the most effective and why?[73]

Approaches to monitoring vary widely. Effectiveness depends on the anticipated types of attacks the system is expected to defend against, the origins of the attacks, the types of assets, and the level of concern for various types of threats (inability to conduct business, damage to reputation, loss of trust, defacement, theft of IP address, dissemination of critical information, unauthorized access, insider policy violations, etc.). In addition, company policy, individual responsibilities, and the purpose/use of the monitored systems/segment determine how threats are identified and categorized and how priorities are established. By addressing these concerns, the institution can determine where the monitoring needs to be performed, what needs to be monitored, and how to interpret events.

[71] Galaxy Computer Services Inc.

[72] Input contributed by Galaxy Computer Services Inc.

[73] In these discussions of the types of monitoring approaches and the three types of detection technologies, many excerpts are from an interview conducted on Jan. 31, 2002, with Linda McCarthy, Vice President of Systems Engineering at Recourse Technologies.

Monitors are categorized based on their placement (network-based or host-based) and the types of sensors they employ (traffic monitors, log file analyzers, system call tracers, integrity checkers, policy monitors, etc.). Regardless of the placement, all monitoring systems are based on three basic detection technologies for identifying security events: signature-based detection, statistical anomaly detection, and protocol anomaly detection.

Signature-based detection is the simplest and most common form of detection. This technology relies on prior knowledge of how an attack exploits a known vulnerability. Signature-based systems monitor a flow of data, looking for previously identified signatures (patterns) of attacks. These systems require continuous updates to ensure that they contain signatures for recent attacks. Because of their architecture, signature-based systems fail to detect new attacks and minor variations on old attacks. Fragroute, a new tool for manipulating packets of data that travel over the Internet, is helpful in detection, but it is also helpful to intruders. Its capability to illuminate weaknesses in a network's security can aid a system administrator in protecting the network. But it has the potential to allow attackers to camouflage malicious programs just enough to bypass many intrusion-detection systems and firewalls. It has several techniques to fool the signature-based recognition systems used by many intrusion-detection systems and firewalls. The program exploits several ways of inserting specific data into a sequence of information in order to fool detection programs.

Statistical anomaly detection models a system’s normal behavior and attempts to identify variations that are indicative of a security event. This is generally accomplished with technologies such as behavior profiling, predictive pattern generation, model-based event analysis, state transition analysis, statistical transaction analysis, data mining, classification, and link and sequence analysis. These systems typically rely on sophisticated artificial intelligence engines, probabilistic reasoning systems, genetic algorithms, or neural networks to build appropriate learning algorithms. The most commonly employed learning algorithms include decision trees, decision rules, linear models, instance-based schemes, numeric prediction techniques, and clustering. Statistical anomaly detection suffers from problems associated with ground-truth reliance, where the system may be rendered useless if it builds a model of “normal” behavior during a period of abnormality. This could occur if the system being modeled has already been compromised. Statistical anomaly detection systems tend to have long learning cycles and problems with high false-positive rates.

Protocol anomaly detection monitors communications for protocol violations or abnormalities in the exchange of information. A protocol is a set of rules that describes the interaction among elements in a computer system. The vast majority of network-based attacks exploit weaknesses in protocol definitions and/or implementations. By modeling a protocol’s normal usage, a protocol anomaly detection sensor can detect these attacks by identifying traffic that does not follow normal protocol behavior. Protocols define a theoretical method for computer systems to interact that is typically described in a standards document such as an Internet Engineering Task Force request for comment. In practice, however, protocols are seldom used in complete accordance with the published standard. As a result, a protocol anomaly detection sensor must model normal usage by taking into account the theoretical definition and the typical usage actually seen in practice. Unlike signature-based systems, protocol anomaly sensors can detect new attacks and do not require continuous updating. In addition, they tend to generate fewer false positives than statistical systems, and they do not suffer from ground-truth reliance or long learning cycles.

Regardless of the method employed, network systems should be monitored at least daily, depending on the amount of traffic allowed into the network by the firewall. It should be pointed out that an IDS is usually considered a post-event security device. Depending on how it is configured, it may be able to notify you that someone has compromised your systems, but it may not be able to stop the compromise. By far the most effective means to stop intruders is to routinely scan your systems for vulnerabilities and to repair them before they can be exploited.

Correlation between firewall and IDS event information is very useful for systems administrators. Ideally, IDS sensors "book-ended" around firewalls will characterize levels of suspicious activity experienced outside, and the internal IDS and firewall logs can depict how many attacks are allowed through the firewall or generated internally by nature of the firewall rule sets.[74] It helps to attempt to build an automated decision matrix engine to do the comparisons on events and data coming through the institution's perimeter controls.

The most cost-effective method for IDS is as follows. First, develop appropriate policies and procedures for use and security. Second, take advantage of the system's inherent security measures that track all system logins and detect intruders. Third, investigate the use of OpenSource software. Fourth, investigate the use of commercial software to fill any voids. One of the most widely used IDSs is an open source system named Snort.[75] On the commercial side, many experts consider RealSecure from Internet Security Systems to be the de facto standard. Cisco Systems Inc., Netranger, Network Flight Recorder, Enterasys Dragon, and Recourse Technologies are also widely used.

F. Virus Scanners

Worms, Trojans (the analogy is to the Trojan horse), and viruses are vehicles for deploying an attack. A virus is a program that can replicate itself by infecting other programs on the same system with copies of itself. Trojans do not replicate or attach themselves to other files. Instead, they are malicious programs that are hidden within another program or file. Once the Trojan file is executed, it can perform malicious activity at will. Virus scanners are critical in the mitigation of these attacks.

Virus scanners should be updated every night. Beginning with an institution's e-mail gateway, every inbound attachment should be scanned for viruses. Fileservers should be set to active scanning mode where they scan every file copied onto them. Desktop scanners that protect the user's PC should also be updated. Data should be tested against standard loads if updates catch anything.

Worms, which are a relatively new phenomenon,[76] use existing security vulnerabilities to gain access to the device. Worms replicate themselves onto other systems via a network connection.[77] Typically, viruses and worms become malicious only when the infected files are accessed or deployed. Most of the time, these vulnerabilities can be eliminated by simply applying patches. The irony here is that someone who is not keeping up-to-date with patches most likely is not keeping up-to-date with virus software either. This human "system" failure can have catastrophic implications for an institution's e-financial network.

[74] E-Global Review, a Predictive Systems product, depicts these occurrences.
[75] Opinion of Rick Fleming, Vice President of Security Operations for Digital Defense Inc.
[76] Worms manipulate networks to spread from computer to computer. This form of malicious code has the ability to enter a network through an open browser. This characteristic is a result of worms' concentrated attacks on servers. In some instances—Code Red, for example—worms have been used to set up back doors into financial networks.
[77] Interview with Linda McCarthy, Vice President of Systems Engineering for Recourse Technologies.

Best Practices to Prevent Outbreaks in Your System Due to Worms and Viruses[78]

1. Understand that mass mailing worms *will* come from someone you know.
2. Do not open unknown e-mail attachments.
3. Do not connect to Web-based e-mail from the country systems—Hotmail, AOL, or Yahoo, for example.
4. Check to see that your antivirus product is updated at least monthly.
5. Use plain text format for e-mail (not HTML) to prevent embedded malicious code.
6. Use Rich Text Format (.rtf) or plain text (.txt) for your documents instead of document (.doc) format.
7. Save a document received as an attachment to disk and scan it before opening.
8. Consider putting an alias as the first entry in your e-mail address book that will send a message to your system administrator or trusted antivirus lab as an early warning mechanism for a mass mailer outbreak.
9. Set Internet Explorer security settings in the Internet Zone to *High.*
10. Set Internet Security settings to disable ActiveX and Active Scripting. Javascipt settings should be placed on disable or prompt.
11. Disable Open and/or Preview panes if implementing Outlook Express.

For System Administrators

1. Filter executable attachments at the e-mail**.**
2. Subscribe to one of the following mail lists to be proactive and to receive early warnings for newly discovered malicious code:
   http://www.cert.org
   LISTSERV@lehugh.edu
   http://www.ntbugtraq.com
   http://www.2600.com
   http://www.nipc.gov
   listserv@netspace.com
   linux-security-request@Redhat.com

Any reliable antivirus software can be set up to check with the manufacturer and to receive pattern and scan engine updates automatically. However, as a general rule, IT staff should check for updates at least weekly. Most virus scanners can detect and effectively eliminate worms and Trojans found in e-mail and downloaded files. Still, some worms may propagate by exploiting vulnerabilities that are not being monitored by the scanning system. The good news is that most of these are discovered once the scanner is updated and a complete rescan is accomplished.

G. Encryption[79]

Cryptography and cryptographic tools sound complex and mysterious. The details of how these tools are constructed and work are intricate, laced both with mathematics and with provable and unprovable properties. The security of some tools can be based firmly upon some intractably difficult mathematical problem. The security of other tools cannot be proven formally, but is trusted as a result of the inability of experts to find and demonstrate any weaknesses in the tools over periods of years. However, what cryptographic tools do and how they are used are very easy to understand. There are only six basic types of cryptographic tools. They are:

[78] This section was provided by Peter Tippet, Chief Technical Officer of Trusecure Inc.

[79] Inputs for this section were contributed by Bill Worley, HP fellow, Chief Scientist for Hewlett Packard, Inc.

1. Symmetric (secret) Key Encryption.
2. Asymmetric (public/private) Key Encryption.
3. One Way Hash Functions.
4. Message Authentication Codes.
5. Digital Signatures.
6. Random Number Generators.

By careful use of these cryptographic tools one seeks to design systems that can provide system security in the face of any of the attacks defined in an associated threat model.

Encryption/Decryption

Symmetric and asymmetric key encryption both are used to scramble data so completely that an attacker lacking the correct "key" is unable to determine the punctuation or control, or video and audio images.

The process of scrambling the data is called "encrypting," or "enciphering" the data. The reverse operation, to unscramble the data back to its original form, is called "decrypting," or "deciphering" the data. The original unscrambled data is called "clear" text, signifying that the meaning of the text is "clear." The scrambled data is called "cipher" text, signifying that the meaning of the data has been obscured.

There are many different algorithms[80] for encrypting and decrypting. All algorithms for encrypting and decrypting take two input arguments, and produce as output the encrypted or decrypted data, respectively. The first input argument is the data to be encrypted or decrypted. The second argument is called the "key." Best practice deems the secrecy of an encryption and decryption scheme to inhere solely in the key. To describe encryption or decryption, we shall write:

Cipher-text = Encrypt(Clear-text, Encrypt-key);

Clear-text = Decrypt(Cipher-text, Decrypt-key);content or meaning of the original unscrambled data. The data itself can be any form of digitized information — letters or numbers of any language, special symbols for words.

The principle that secrecy must inhere solely in the key was first articulated in January and February 1883, in a book entitled *La Cryptographie Militaire*, first published as two installments in the *Journal des Sciences Militaries*. The author, born Jean-Guillaume-Hubert-Victor-François-Alexandre-Auguste Kerckhoffs[81] on 19 January 1835, and the same publisher later in 1883 reissued Kerkhoffs's book as a paperback.[82] Kerkhoffs's 64-page book, widely regarded as one of the most famous, concise, and important books on cryptology, remains a guiding influence to this day.

Symmetric Key Encryption

For symmetric key encryption, both the sender and receiver of encrypted messages employ the same secret key. That is, in the formulas: Encrypt-key = Decrypt-key. For example, if the secret key were the word "applesauce", the equations would be:

[80] Also called "Ciphers."

[81] He later shortened his name to simply: Auguste Kerkhoffs.

[82] David Kahn, *The CODE-BREAKERS*, Scribner, 1967, 1996.

Cipher-text = Encrypt(Clear-text, “applesauce”);

Clear-text = Decrypt(Cipher-text, “applesauce”).

Of course, the keys that actually would be used by systems would be nothing so simple as an English language word. Rather, they would be long, randomized bit strings. The actual lengths of keys are determined by the specific symmetric cipher. The actual binary values of keys are generated by processes that are constructed to insure that the values are sufficiently random and unpredictable. Symmetric ciphers are of two basic types: “block” and “stream.” A block cipher encrypts or decrypts a single, fixed-size block at a time. The size of a block is defined by the specific block cipher. A streaming cipher generates a sequence of binary values that are XOR’ed[83] with the Clear-text or Cipher-text.[84] The size of each binary value generated by a stream cipher is defined by the specific stream cipher.

The names, key lengths, and block/stream-element lengths of some important symmetric ciphers are shown in the table below. The newest of the symmetric ciphers is the Advanced Encryption Standard (AES). On 6 December 2001 it was approved as FIPS 197, effective 26 May 2002. The cipher was selected on 2 October 2000 as the winner of an extensive cipher design and evaluation competition run by the National Institute of Standards and Technology (NIST).[85]

Cipher Name	*Acronym*	*Key Length (bits)*	*Data Length (bits)*	*Type*
Data Encryption Standard FIPS 46	DES	56	64	Block
Data Encryption Standard (export restricted)	DES 40	40	64	Block
Triple DES	3DES	3 x 56, ~112 [86]	64	Block
Advanced Encryption Standard FIPS 197	AES	128, 192, 256	128	Block
Rivest Cipher 4	RC4	Variable	Byte Stream	Stream

Symmetric key ciphers execute at high speed. On grounds of security and performance, symmetric ciphers rank very highly. For any confidential, high-volume data interchange, symmetric key cryptography is the preferred choice, simply because of its high performance. The hard problem for deploying symmetric cryptography is the distribution and management of secret keys, the aptly named “Key Distribution Problem.”

The Key Distribution Problem, i.e. the definition of a process securely to distribute unique secret keys throughout a network, to every pair of persons or programs that needs to communicate in confidence, is very complex. Entire systems have been built solely to perform this function. The “Kerberos”[87] system developed by MIT, based upon the work of Needham and Schroeder, dedicates entire servers simply to the task of being trusted third parties for distribution of secret keys. We shall not go into the details of Kerberos. We simply observe that it is a workable symmetric key distribution system, that it is employed in Microsoft’s security strategy, that it has

[83] XOR is the exclusive OR logical operation: (0 XOR 1) = (1 XOR 0) = 1; (0 XOR 0) = (1 XOR 1) = 0;

[84] The same stream value can be used both for encryption and decryption. This works because XOR operands can be ordered and grouped arbitrarily, the XOR of any value with itself is zero, and the XOR of zero with any value yields the value itself. In symbolic terms:
Cipher-text = Clear-text XOR str-val;
Clear-text = Cipher-text XOR str-val = (Clear-text XOR str-val) XOR str-val =
Clear-text XOR (str-val XOR str-val) = Cleart-text XOR ‘zero’ = Clear-text.

[85] The finalist ciphers were named MARS, RC-6, Rijndael, Serpent, and Twofish. The selected winner was Rijndael (pronounced: “Rain-doll”), named after portions of the names of the two inventors.

[86] Three 56-bit keys are used, but a known attack reduces the effective key strength to 112 bits.

[87] “Kerberos” is the name of the three-headed dog that guards the gates of Hades.

evolved to its current Kerberos Version 5, and that in many cases, the need for a key distribution system such as Kerberos entirely can be avoided through the use of asymmetric key encryption, which is the subject of the next section.

Asymmetric Key Encryption

For asymmetric key encryption, the key used to encrypt data is a different key from that used to decrypt data. Unlike symmetric key encryption, which uses the same secret key both for encryption and decryption, asymmetric key encryption uses two different keys. Why is this important? It is important because only one of the keys needs to be kept private (i.e. secret). The other key can be widely known, i.e. can be made public. It is for this reason that asymmetric key encryption popularly is called "public/private" key encryption. Asymmetric ciphers greatly facilitate problems of key distribution.

Symbolically:

Cipher-text = Encrypt(Plain-text, public-key);

Plain-text = Decrypt(Cipher-text, private-key);

One can appreciate the power of having two keys by considering how one orders items over the Internet, such as books from Amazon.com. The ordering process used by such websites uses a protocol called SSL,[88] to assure that a secure session is established between the customer's computer and the website. An SSL session begins with a "handshake" protocol. The customer's computer sends a "hello" message to the Amazon web server, and web server's reply includes a Certificate containing an Amazon public key. The customer's computer checks that the Certificate is valid,[89] and then uses Amazon's public key to encrypt data that both the user's computer and the web server will use to construct a symmetric key for the session. Only the Amazon web server has the private key needed to construct the symmetric session key. After some further checking, the session continues using the symmetric key to encrypt/decrypt messages. The order information—credit card number, shipping address, gift-wrapping, greeting message, and items ordered—then can be sent confidentially to Amazon.

The most widely used asymmetric cipher is called "RSA," an acronym composed of the first letters of the last names of its inventors: Ron Rivest, Adi Shamir, and Leonard Adleman, who first published their work in the summer of 1977.[90] Patent protection for the algorithm expired in September 2000, and it now is in the public domain. RSA operates upon very large integers modulo the product of two secret prime numbers. RSA public and private keys each are pairs of such large integers. For good security today, 1024 to 2048 bit integers are used, although some applications continue to use 512 bit integers. The cryptographic strength of RSA derives from the difficulty in factoring the two secret primes given only their product.

More recent work has shown that finite groups of points lying on an "elliptic curve" provide a perhaps stronger base than RSA for asymmetric cryptography. The security of the schemes based upon elliptic curves derives from the difficulty in solving the discrete log problem over elliptic

[88] "Secure Sockets Layer," deployed in nearly every web browser and web server throughout the Internet. For details, see: Eric Rescorla, *SSL and TLS*, Addison-Wesley, 2001.

[89] This is done by validating the digital signature of the Certificate. The public key for this validation is that of the Authority issuing the Certificate, which normally is stored within the browser in the customer's computer.

[90] It was later learned that Clifford Cocks, of the British General Communication Headquarters (GCHQ), independently in late 1973 had discovered the same algorithm.

curves.[91] The keys needed for computations are shorter than those needed for RSA. The elliptic curve key equivalent in strength to a 1024-bit RSA key, for example, would be less than 200 bits in length. The computations, however, are more complex than those for RSA.
The primary difficulty for asymmetric encryption is that it requires much, much more computation. In other words, it is a lot slower than symmetric encryption. For example, AES on the Itanium 2 microprocessor operates at approximately one bit per cycle. For a 1024-bit value, AES would require about 1024 cycles to encrypt or decrypt. In contrast, on the same microprocessor, 1024-bit RSA would require about 20,000 cycles for a public key encryption, and over 250,000 cycles for a private key decryption.

Because of the relative strengths and weaknesses of symmetric and asymmetric cryptography, a common practice is to use asymmetric key cryptography for key distribution, and symmetric key cryptography for the bulk of the transferred data. This is what is done within the SSL protocol—asymmetric cryptography is used to establish a newly created symmetric key, which then is used for the data transfers within the SSL session.

One-Way Hash Functions

One-way hash functions[92] are algorithms that transform an arbitrary input bit stream into a fixed-size result. They are called "One-Way" because while it is straightforward to perform the hash function, it is provably impossible to compute the inverse of the hash function—that is, to compute the input bit stream given only the fixed-size hash function output.

Further, secure hash functions are designed so that, to the extent possible, the value of every bit of the output of the function is affected by the value of every bit of the input, and that for a given hash output value it is a prohibitively difficult task to find another input bit stream that would produce the identical hash function output. These properties of secure hash functions permit one reasonably to use the hash output value as an identifying "signature" or "digest"[93] for the input bit stream. In Bruce Schneier's parlance, hash function outputs are like "digital fingerprints" of the input bit stream.[94] Others have characterized hash outputs as "unforgeable check sums."

Secure hash functions typically operate in two phases. First, the arbitrary-length input bit stream is padded in a specific manner to a multiple of a fixed block size. Usually, the bit length of the entire input stream appears as the final 64-bit value in the last 64-bit double word of the final padded block. Second, the padded input is processed a block at a time by the hash function, using output values from each block as initial input values to the processing of the next block. The combined result after processing the final block is the output of the hash function.

Symbolically:

```
Message Digest = Hash-Function( Pad-Function( Input-Bits ) );
```

The two most widely used secure hash functions are "SHA-1" and "MD5". SHA-1 is the U.S. Governments' standard hash function. The acronym "SHA" stands for "Secure Hash Function"; the algorithm is documented in the publication: "Secure Hash Standard", FIPS PUB 180-1. The output of SHA-1 is 160 bits in length. RIPEMD-160 is a similar European algorithm. The recently defined SHA-256, SHA-384, and SHA-512 functions produce hash outputs of 256, 384,

[91] For details, see: Blake, Seroussi, Smart, *Elliptic Curves in Cryptography*, Cambridge Univ. Press, 1999.
[92] Also called "Secure Hash Functions."
[93] For this reason, the output of a secure hash function also is called a "Message Digest."
[94] B. Schneier, *Secrets & Lies – Digital Security in a Networked World*, John Wiley & Sons, 2000.

and 256 bits respectively. MD5 stands for “Message Digest 5”, and was invented by Ron Rivest. MD5 produces a 128-bit hash output. Although still in use, MD5 has shown some cracks and now is seldom used for anything new.
Hash functions are one of the most versatile and widely employed cryptographic tools. They are used in nearly every Internet protocol. Whenever it is necessary logically to associate a number of elements, the common practice is to concatenate the byte-strings representing the elements and hash the total result. This produces a digital fingerprint reflecting every element as well as the entire association. Hash values of documents, data structures, and code images are crucial for digital signatures, which are discussed below.

Message Authentication Codes

Message Authentication Codes (MACs) are designed to protect the authentication and integrity of messages and data rather than to protect their confidentiality. A MAC is a value appended to a message or data that permits one to assure that the content of the message or data has not been modified in any way.

MACs, like symmetric key encryption, use a secret key. If it is not important to protect the content of a message, but only to insure its integrity, the secret key is used to compute a MAC, which then is appended to the message. Any party who knows the secret key can recalculate the MAC from the received message, and compare the newly calculated value with the MAC appended to the received message. If the newly computed and transmitted values match, the message has been transmitted correctly. MACs are used in secure IP (IPSec) to assure that packet contents have not been modified during transmission. They also are used in transfer protocols between banks to authenticate messages. MACs also can be attached to data stored in files or databases. In all cases any one knowing the secret key can verify that the data over which the MAC was calculated has not been altered. Computations of MACs utilize symmetric key encryption over the data, or sometimes over secure hashes of the data.

For confidential storage or transmission of data, both MACs and encryption must be employed. For example, first a MAC is calculated using a secret key and then appended to the block of data; second, the data and its appended MAC are encrypted using a second secret key. This assures both that the data remains unintelligible to unauthorized parties and that its integrity remains intact. It is becoming generally accepted that this combined protection is the proper way to secure network transmissions.

Digital Signatures (PKI)

Digital signatures are similar to MACs in providing a guarantee of the integrity of stored or transmitted data. But digital signatures differ from MACs with respect to the secrets employed and the distribution of those secrets.

For MACs there is a non-empty set of persons who know the secret key used to compute the MAC. Any of the people in the set can verify a MAC computed with the secret key. At the same time, any of these persons also can compute a MAC for different messages or data, or can forge a MAC for an altered message or data. All the persons who know the secret must trust each other—both to keep the secret and to employ it properly.

MACs work well when there is but a single individual who holds the secret. This person can use the secret to protect his or her data and to assure its integrity. It also works well for sets of two persons. These folks trust each other not to reveal the secret, and can send messages back and forth knowing that the integrity of each message has not been compromised. These models fit many important situations.

Digital signatures are an integrity protection mechanism where there is but a single party who is capable of constructing the signature, but everyone then is able to verify the signature. In this model, one specified party, the sole holder of the enabling secret, becomes responsible for computing and appending the signature to the data. This is called “signing” the data. After the data has been “signed”, anyone can verify that the specified party in fact did compute the signature. This model fits many important commercial and practical situations.

Digital signatures are computed by using a remarkable property of asymmetric key cryptography. Namely, it also works in reverse. In the previous section on asymmetric key cryptography we wrote symbolically:

Cipher-text = Encrypt(Plain-text, public-key);
Plain-text = Decrypt(Cipher-text, private-key).

But it turns out that the keys can be used in the reverse order, namely:

Cipher-text = Encrypt(Plain-text, private-key);
Plain-text = Decrypt(Cipher-text, public-key).

In the first case the owner of the secret key does the decrypting, and anyone can do the encrypting. In the latter case, the owner of the secret does the encrypting, and anyone can do the decrypting.

Digital signatures often are used by themselves, rather than in combination with encryption designed to protect confidentiality. Usually, it is not important that the data contents be confidential, but it is extremely important that the data be accurate. Digital signatures are not computed by encrypting all of the data with the private key, but rather are computed by encrypting a secure hash of the data by the private key. In effect, we first take a digital fingerprint of the data, and then encrypt the digital fingerprint with the private key. Symbolically:

Digital-Signature = Encrypt(Hash-Function(Input-Bits), private-key);
Digital-Fingerprint = Decrypt(Digital-Signature, public-key).

Digital signatures also are verified in two steps. First the digital fingerprint of the data contents is recomputed using the hash function. Then the appended digital signature is decrypted using the *public* key. If the recomputed digital fingerprint and the decrypted digital signature match, the signature has been verified. The mathematics look very easy, but some delicacy is required to construct correct protocols around them. For example, if a digital signature is appended to a letter not including the identification of the intended recipients, an attacker may combine the body of the letter with a fictitious list of recipients and falsely claim that the original signer sent the letter to the fictitious list. Basically, digital signatures are used to protect the integrity of a collection of data. They often therefore are used as analogs of written signatures for documenting transactions, with and without the participation of trusted third parties. However, they are not exact analogs of written signatures, and certain niceties carefully must be observed.

There are several commonly used digital signature algorithms. RSA private key encryption applied to SHA-1 message digests of the signed material is the most common. The US government has defined as its digital signature standard a more complex algorithm known as

DSA, which also operates upon an SHA-1 message digest.[95] Signature algorithms invented by Taher ElGamal[96] and based upon elliptic curves also are employed.

Random Number Generators

Random numbers are employed throughout cryptographic algorithms and protocols. They are used for keys, challenge values, pre-hashing appendages for passwords,[97] etc. Hardware devices based upon some form of physical randomness are beginning to appear. The problem with such hardware devices, of course, is testing them to assure they're operating correctly.

Solely computational means for generating truly random numbers do not exist. A favorite quote of John Von Neumann's, cited by Bruce Schneier, is "Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin." Fortunately, computational means do exist for computing numbers that are sufficiently unpredictable that they can be used in lieu of truly random numbers. Such numbers are called "Pseudo-Random Numbers" (PRNs). Some of the pseudo-random number generation methods employ values obtained by physical measurements of random events in a computer system, such as typing rates, arbitrary mouse motions, arrival times of I/O interrupts, etc. Others are based upon symmetric cryptography[98] or the difficulty of hard mathematical problems such as the factoring problem.[99] Pseudo random number generators (PRNGs) that produce sufficiently unpredictable values are called "Cryptographically Strong Pseudo Random Number Generators' (CSPRNGs).

A trusted system must include one or more CSPRNGs. Furthermore, it must provide means for storing any internal state needed by the CSPRNGs in a location that cannot be accessed as the result of any system penetration. The latter requirement is particularly difficult. In UNIX, Linux, and NT systems any hacker penetration into privileged mode permits code to examine all of physical memory, including stored CSPRNG state. A paper by Adi Shamir and Nicko van Someren[100] described the algorithms that permit detection of likely cryptographic keys and other random-looking materials in memory. Such results then can be forwarded over the network to an attacker's machine, and analyzed there for use in attacks against system data.

One of the easiest points of attack against encrypted data is its underlying pseudo-random number generation. The security of SSL V2, for example, was broken by Goldberg and Wagner[101] when, from a code listing, they discovered that the PRN was seeded with the time of day and process ID. These values were sufficiently predictable that the entire space could be searched in about an hour's computer time (less on today's computers). A Security Strategy must assure that all such avenues of attack are eliminated completely.[102]

[95] Digital Signature Standard, FIPS PUB 186. DSA always was in the public domain.

[96] T. ElGamal, "A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms, " *IEEE Transactions on Information Theory*, volume, IT-31, 1985.

[97] Called "salt."

[98] E.g. ANSI X9.17

[99] E.g. the Blum-Blum-Shub (BBS) algorithm, L. Blum, M. Blum, and M. Shub, "A Simple, Unpredictable Pseudo Random Generator," *SIAM Journal on Computing*, volume 15, no. 2, 1986.

[100] Adi Shamir and Nicko van Someren, "Playing hide and seek with stored keys." 22 September 1998.

[101] I. Goldberg and D. Wagner, "Randomness and the Netscape Browser," *Dr. Dobbs Journal*, January 1996.

[102] The most common way encryption is compromised is that the password or phrase used is easy to guess, and the hacker simply lucks out and uses the correct password. The other way is through a fundamental flaw in the algorithm being used to encrypt the information, which allows for a successful attack.

H. Securing Servers

Best Practices

1. There should be no remote control and/or administration of host systems from other than the host system console unless explicit measures such as Secure Shell (SSH),Secure Socket Layer (SSL), or Virtual Private Network (VPN) have been made to secure both the network and host system environments. This includes, but is not limited to, system administrator access via modems, gateways, bridges, routers, switches, protocol converters, and other terminal-to-host connections.
2. All security functions and software changes/additions should be made only by authorized server administrator personnel; all changes should be documented.
3. The general user should not have access to the system console; system administrator logons should be done at the system console or any other end-user device if such capability exists.
4. There should be no access to the host system or to any system resources following diskette boot for those systems without physical security.
5. There should be no peer-to-peer communications between general-purpose end-user workstations unless authorized by the system administrator.
6. There should be only one user ID per user on a single system. There also should be only one user making use of any user ID, except as documented and approved in writing.
7. There should be no unauthorized or unsupervised use of traffic monitors/utilities (e.g., Data General Sniffer), recorders, or other customer premise equipment.
8. The host server and console should be physically and logically secured from unauthorized access and use. Maintenance performed on a host server by anyone other than authorized personnel must receive documented authorization from a system administrator.
9. Workstation sessions should be suspended and/or terminated after a predetermined period of inactivity, if technically possible.
10. In cases where session slippage occurs, or in instances where service requests require significant changes of access level privileges, user reauthentication should be required.
11. Successful logons should display the date and time (duration) of the end-user's last session. This is especially useful for super-user accounts.
12. The integrity of data should be maintained by using transaction locks on all shared data for both data files and databases.
13. Administrators should have the capability to constantly scan for viruses at the host server level, the workstation level, and so on.

Microsoft IIS Web Servers

The vulnerabilities inherent in remote data service (RDS) gave crackers the ability to exploit ODBC, which allows the execution of DOS commands within queries that result in a cracker's ability to run arbitrary commands.[103] The RDS vulnerability can be fixed by referring to the following URL: http://support.microsoft.com/support/kb/articles/q184/3/75.asp. Microsoft IIS Web servers continue to support many different programs such as HTR, STR, and IDC. These programs have a variety of associated vulnerabilities, "primarily the exposure of source to scripts." Also, "RDS, DOS commands, and HTR are all part of the default installation of IIS

[103] The source for the information and quotes in this paragraph is "Secure Your Microsoft IIS Web Servers," a presentation of Trusecure on April 18, 2001.

servers." In its effort to make its software more accessible to its technical support and to allow for easy entry into any entity's system to repair software problems, Microsoft created back doors or points of entry that are not published. However, this feature provides easy access for criminals as well as trained technicians. To mitigate the risk associated with these servers, it is necessary to remove all programs except those that are implemented exclusively within your own sites.

Microsoft Security Bulletins can be searched at: http:///www.microsoft.com/technet/security/current.asp. The tragedy of September 11 made it clear that this host should not be located at headquarters. Instead, a remote secure facility should house all data backup servers.

The recent release of Security Assertion Markup Language (SAML) is supposed to "secure" all Windows 2000 operating systems. Security can never be guaranteed. In reality SAML is merely a "box full of patches" that will give users the illusion of safety when in fact it is merely a box full of Band-Aids. Silver bullet solutions do not exist.

I. Penetration Testing

It is not easy for banks and financial institutions to safeguard their networks because new vulnerabilities are discovered daily, and their fixes/patches must be diligently applied to all systems. New connections to the Internet, modems, and virtual private networks (VPNs) create a multitude of new access points to a network whose risk is defined by its weakest link.

Penetration testing entails obtaining knowledge of existing vulnerabilities of a computer system or network and using that knowledge to attempt to gain access to resources on the computer or network while bypassing normal authentication barriers. It may also include exploiting vulnerabilities to gain increased authorization—for instance, to go from regular user to super-user. Penetration testing is good only on the day it was done (this is true for *all* security testing). Penetration testing is an excellent way of testing installed security measures, policies and procedures, and the effectiveness of a company's end-user security training programs. First, a company will be able to tell if security measures such as firewall and IDS systems are functioning properly and what skill level is required to circumvent them. Second, a company will gain insight into whether established policies and procedures allow its staff to detect and react to an intrusion properly. And third, a company can determine if additional training is required for its end-users.

Penetration testing should be performed at least annually, and more often if the system is subject to frequent application or operating system updates. Once a penetration test has been performed, ongoing vulnerability assessment should be performed to address newly discovered exploits. The frequency of the vulnerability assessment should be determined on the level of risk an organization is willing to accept. Given the speed at which new vulnerabilities and exploits are discovered, a vulnerability assessment should be performed semiannually and in many cases quarterly no matter what level of risk one is willing to accept.

Proper Systems Administration

The following is a list of typical administrative failures within financial institutions and corporations.

Top 20 red flags

1. Lack of training and expertise of administrators.
2. No time for or interest in reviewing log files.
3. No time for or interest in hardening machines.
4. Deployment of new technology without security peer review.
5. Failure to install software patches that fix security flaws.
6. Lack of strict requirements to use strong passwords.
7. Removal of security mechanisms because they cause user inconvenience. Restoration of systems from backups and failure to reload any patches that were previously installed.
8. Failure to remove administrative-level accounts that were added temporarily for service personnel.
9. Failure to install or use available security mechanisms such as password policy enforcement or system event logging.
10. Lack of daily audit of network logs for suspicious activity.
11. Setting up computer systems using the software defaults. These default settings are designed to get the system up and running with the least interference and are often very insecure.
12. Failure to perform routine backups of systems and then test those backups for viability.
13. Failure to properly install and update virus protection software.
14. Sharing administrative accounts and passwords over multiple systems.
15. Primary reliance on a firewall or public key infrastructure (PKI) system for security.
16. Use of SNMP, telnetd, ftpd, mail, rpc, rservices, or other unencrypted protocols for managing systems.
17. Assignment of passwords to users over the phone.
18. Failure to educate users about security problems and what to do when they see a potential problem.
19. Poorly written and implemented policies and procedures.
20. Improper documentation.

Best Practices

1. Network administrators should be responsible for installing and verifying patches and updates to operating systems weekly.
2. Onsite trained security staff should be present 24/7.
3. Employees should be required to use robust passwords (long in length; mix of letters, numbers, and symbols), which should be changed monthly.
4. Computer monitors should not be visible to anyone who is not an employee of the institution.
5. Network administrators should implement a profile procedure to process employees transferring to another office in the bank, termination of employees, and changes to an employee’s level of access within the bank's systems.
6. Those who are responsible for large value transfers should utilize biometric identifiers as their password.
7. Backups should be maintained of all critical material that is stored in a different location.
8. An incident response capability and a plan that ensures continuity of operations and recovery from security breaches should be in place.
9. Strong authentication—preferably biometrics, smart cards, and cryptography—should be exercised for large value transfers.
10. Firewalls and intrusion detection systems should be installed.

11. Penetration testing/auditing should be performed on all of the institution's systems.
12. A login banner should be displayed stating that the system is only for authorized use and is subject to monitoring.
13. Patches must be updated weekly to both servers and remote access machines. See http://www.microsoft.com/technet/security/current.asp
14. Critical operations should have two-person controls.
15. A security policy should be developed that mandates training for non-IT staff vis-a-vis an incident response plan and that prohibits instant messaging, voice-over IP, and wireless local area network (WLAN) installation without appropriate authorization and securitization.

Adequate Backup Procedures

All system and user files should be backed up on a regular basis. Develop a plan that is broad enough to cover all the workstations and servers you have deployed. If you have regularly created checksums for all fields and have securely stored these checksums, you can plan to restore files from trusted backups against which the checksums are calculated. Backups should be centrally administered, with data copied from workstations via networks. Encryption tools can be used to protect data passing from a user's workstation to the central backup host. Backups must be made at least daily. The minimum requirement in most organizations is to perform a full backup weekly and incremental backups every day. At least once a month the backup media should be verified by doing a restore to a test server to see that the data is actually being backed up accurately.

J. Policy Management Software[104]

Given that good e-security is a combination of people, processes and technology, Banks should be implementing a policy approach that is governed by system wherein the policy and enforcement are dynamic. Bank policy vis-à-vis computer usage necessitates enforcement by a software program. The verbal policy dimension should be translated into machine code. This method of policy enforcement mitigates the insider threat dimension both premeditated and accidental. Once policy is built and subsequently amended, users must be educated and then regulated by a rule-set which is modular not static.[105] This security approach needs to take into account the privacy rights of the user on the system. Users should not only be identified once an alarm (policy violation) has occurred.

K. Incident Response[106]

The ability to react quickly to security incidents is an essential part of an overall security plan. An organization's ability to operate will depend on its ability to provide timely information to its clients in the form of electronic data.

It is also essential to categorize information. Information from critical systems will surely receive a more direct and focused response than, for example, electronic information stored for office supplies. An organization needs the ability to react to and recover from security incidents as they arise with an effective and coordinated response, which in turn will minimize the cost and damage to the organization's infrastructure and to its image within the banking industry.

[104] Contributed by Keith Schwalm, Director of Infrastructure Protection, The President's Critical Infrastructure Protection Board.

[105] Computer Associates, Tumbleweed and Polivec vend effective policy management software.

[106] Contributed by Simon Martinez Sr., System Analyst at Integrated Management Services Inc.

A security incident can be defined as an event that changes the security posture of an organization or circumvents security polices developed to prevent financial loss and the destruction, theft, or loss of proprietary information. It is characterized by unusual activity that causes the organization to investigate because the activity cannot be explained through normal operations.

Some possible classifications for security incidents are these:

- Virus attacks (unable to clean, rename, or delete);
- Denial of service attacks;
- IDS alert notifications (false positives possible);
- Automated scanning tools.

Banking organizations must share in the responsibility of coordinating their response efforts with those of other financial institutions. Networking in a trusted environment and sharing incident information and detection/response techniques can be important to all of these organizations in identifying and correcting weaknesses. Gathering intelligence information from all sources is a critical part of information infrastructure protection. Having an information-sharing network in place can also help government agencies alert other agencies to potential and/or actual threats directed at the critical information infrastructure of nations.

Incident response within any organization must begin with management. Management is responsible for providing the support, tools, personnel, and financial backing needed to ensure the success of the incident response team. An incident response team must be perceived well by all concerned. Security awareness training and briefings for senior management are key components of a successful deployment of an incident response team.

Sources of information for building a Computer Security Incident Response Team are as follows:

- http://www.securityunit.com/pubs/index.htm SecurityUnit, Inc.
- http://www.cert.org/csirts/ CERT Coordination Center
- http://www.cert.org/training/2002/creating_csirt.html Creating a Computer Security Incident Response Team

To maximize the full potential of the team, members must be available 24/7. Attack can come at any hour. Intrusion Detection Systems (IDSs), network- and host-based, are playing a more critical role in identifying attacks and unusual activity. Alerts from such systems are generated at all hours of the day. An incident response team allows an organization to respond to alerts generated by automated systems 24/7. Monitoring systems and reviewing security alert information submitted by vendors is an important part of an incident response team's proactive duty. IDS systems, however, do not provide a complete solution to identifying and responding to incidents. An overall security plan is needed to ensure overall protection that would include an incident response mechanism.

An incident response team must also develop procedures. Clear definitions of each type of incident will enable members to react quickly and effectively. Procedures must detail the steps team members should take when alerted to an incident. Included within the procedures must be clearly defined investigative goals to be achieved before an incident can be closed. The team should also list and post contact information of key personnel and management to notify.

The team may need to contact other organizations to assist in the investigation. The bank must develop a policy that clearly describes the bank's position on the disclosure of incident information to the banking community as well as outside organizations such as the National

Infrastructure Protection Center (NIPC), the Computer Emergency Response Team (CERT), FedCIRC, and commercial incident response teams. Bank organizations may designate an individual (job function) to coordinate the exchange of information. All team members must sign a nondisclosure form.

Tracking of security incidents can become a full-time job, because all incidents must be tracked. Incidents may remain open from a few hours to a few months, or even longer in some cases. The incident (case) record must contain all communications relating to the incident from the time it is opened to the time of closure. Depending on the type of incident, careful consideration should be given to collecting any data that may be relevant to the incident. Response team members should receive professional training in handling and collecting evidence (system logs and backup tapes) in case such evidence needs to be used in a court of law.

The Incident Response Plan

An incident response plan (IRP) is the primary document an organization uses to establish how it will identify, respond to, correct, and recover from a computer security incident. Every organization should have an IRP and should test it periodically.
All employees should be trained in the correct procedures to undertake in the event of a computer incident. An incident response plan might make the following points:[107]

1. The institution’s security department, legal department, and public relations department should jointly develop and implement an incident response policy.[108]
2. You should contact incident response agencies responsible for your site. For example:
   http://www.nipc.gov (National Infrastructure Protection Center)
   http://www.cert.org (Computer Emergency Response Team)
   http://www.cert.dfn.de/eng/csir/europe/certs.html (European Computer Emergency Response Teams)
   http://www.ectaskforce.org/Regional_Locations.htm (Electronic Crimes Taskforce)
   http://www.first.org/team-info/ FIRST (Forum of Incident Response and Security Teams)
3. Make communication via an out-of-band method (e.g., a phone call) to ensure that intruders do not intercept information.[109]
4. Document your actions (e.g., phone calls made, files modified, systems jobs that were stopped).[110]
5. Make copies of files the intruders may have left untouched (e.g., malicious code, log files) and store them offline.[111]
6. Contact law enforcement officials. To ensure proper reporting during an e-security incident, please refer to the following URL: http://www.nipc.gov/incident/cirr.htm.

A brief discussion of each of these areas of the plan follows.

The term *incident determination*[112] describes the process used to define events as an “incident” and explains how each type of incident should be handled. Indicators that signal an incident may

[107] Provided by Bob Weaver of the New York Electronic Crimes Taskforce.
[108] Ibid.
[109] National Infrastructure Protection Center’s advice.
[110] Ibid.
[111] Ibid.

be categorized as Possible, Probable, and Definite. In addition, if there is a predefined set of conditions that constitute an incident, the incident will be handled in accordance with the plan once those conditions are determined to exist. Arguably, this is the most important section of the plan because it identifies certain situations or conditions and sets out in detail how to respond.

The term *incident notification* describes the procedures to be used in notifying the computer user population once an incident has been confirmed. This section of the plan identifies those who must be notified in the event of an incident and provides critical contact information and contact procedures. These are some elements that may be included in this portion of the plan:

- Internal components of the organization, including management, operations, security, public relations, and the general employee/user population
- Computer security incident response organizations
- Affected partners or other integrated entities
- The organization's insurer, if the organization is insured through an e-risk policy
- Law enforcement at the local, state, and federal levels
- News media and other public relations components

Incident containment is the third area of the plan. It addresses the measures that must be taken to halt/mitigate the effects of the incident and to regain control of the affected networks, systems, and related components.

Damage assessment is a critical step once containment has been achieved. This phase assesses the damage that has been inflicted on the institution’s assets. It should determine the scope of damage, the duration of the incident, the cause of the incident, and the identification of the responsible party.

Incident recovery is the next key element of the plan. It requires a comprehensive approach to returning networks and systems to normal operations. The following are vital activities that must be addressed during the recovery phase of incident response. First, the vulnerability that allowed the incident to occur should be mitigated. If removal of the vulnerability is not possible, then safeguards should be improved to mitigate additional damage. Second, all out-of-date detection mechanisms should be updated. Software, data, and services should be restored in accordance with approved procedures. System and network activity should be followed closely after restoration for any signs of a follow-up attack. Finally, during recovery, efforts must be made to restore confidence in the organization by its users, partners, and clients.

Reflection, or lessons learned, is the final section of the plan. This requires the institution to conduct a postmortem of the incident. The institution reviews what occurred, how it occurred, how it could have been prevented, and what changes need to be made to ensure that such an event does not occur again. In addition, related plans should be reviewed to determine whether any changes in those plans are indicated also.

[112] In the next paragraphs, the terms used for the plan, and the descriptions of the terms, were contributed by Galaxy Computer Services Inc.

Box 2. Survivable System Development

Survivability analysis or business continuity helps to identify the essential functions or assets in the institution that must survive in the event of an attack or system failure. The delivery of essential services and the preservation of essential assets during a compromise, and the timely recovery of full services and assets following attack are among these functions. The organizational integration phenomenon that typifies the modern banking community is accompanied by elevated risks of intrusion and compromise.

It is essential to determine what elements in the institution’s IT infrastructure are absolutely mission-critical—that is, what elements must be up and running within a certain time in order for business to continue. From this point forward one must envisage various compromises to the system so that contingency plans are in effect to cover all potential threats. The following sources on survivability will assist network architects in determining the impact of certain accepted risks.

The Carnegie Mellon Software Engineering Institute Network Systems Survivability Program uses incident data collected by the CERT as a basis for the institute's survivability research and for trend identification and the prediction of future problems.
http://www.cert.org/nav/index_purple.html

The European Dependability Initiative (http://www.cordis.lu/esprit/src/stdepnd.htm) represents a major research effort within the European Union to address the critical infrastructure protection and survivability efforts of the member nations. There are plans for joint U.S./EU cooperation.

The National Infrastructure Protection Center (NIPC) is the U.S. government’s focal point for threat assessment, warning, investigations, and response to attacks (http://www.nipc.gov).

Source: Joe McCleod of CERT

Annex II. Authentication and Nonrepudiation

The Internet economy is built on information. In this economy, time is money, and information is valuable. The value of e-finance is defined, in part, by technology's ability to move information and to affect markets quickly. The underlying assumption is that moving information is reliable. Reliability is based, in part, on constructing a system and a process that keeps the percentage of repudiated transactions to a minimum.

In order to construct such a system, transactions must be appropriately authenticated, verified, and authorized. A precursor to this is access controls. Access controls enable a dumb operating system to know whether an individual attempting to enter the system has been granted access. Authentication is the means used to assure the system that the party attempting to engage in an activity is, in fact, the party so designated. Verification is the means used to confirm that the party claiming a certain identify is the right party. Finally, authorization is the means used to determine that the party engaging in a transaction has the requisite authority to access that portion of the system or to engage in that type of activity.

The value of information is based on its reliability and its integrity—whether the party was authorized to access or engage, whether the identity was authenticated, whether there is a risk of nonrepudiation, whether there are any process restrictions for the particular transaction (specifically, whether the rules engine has any access controls), and whether there are any relationship constraints (specifically, whether privacy or confidentiality is protected). The process restriction is an internal risk, and the relationship constraint is a potential legal liability.

However, the value of any information is directly related to the extent to which the information meets these criteria versus the extent to which it needs to meet this criterion. So on a scale of 1 to 10, if the information should be a 10 but the system can only "assure" 5, it has lost at least 50 percent of its value. Thus, security is a value-added proposition and is a major business consideration.

Table 1. Characteristics of Data

	Authentication	*Integrity*	*Reliability*	*Nonrepudiation*	*Process Restrictions*	*Authoritative*	*Relationship Constraints*
Biometrics	▪	▪	▪	▪	▪	▪	▪
PKI	▪			▪	▪	▪	▪
Smart Card	▪				▪		▪
Notary License			▪	▪		▪	
Digital Time Stamp			▪	▪			

As a result of the information age, Web sites that cater to hackers are springing up. The Web site http://astalavista.box.sk represents the growing pantheon of hacker search engines that disseminate critical weaknesses in existing hardware, software, encryption technologies, and wide area networks (WANs). It has shifted from closed architectures to networks—intranets and extranets, local area networks (LANs), WANs, virtual private networks (VPNs), satellite, microwave, and wireless. Distributing information by broadcasting makes it more difficult to protect assets, monitor operations, and respond to problems. Turnover in the workplace from executives to managers—as companies are raided for talent—creates risks, such as theft or destruction of intellectual property and theft of access codes. It is evident that distinguishing one's identity over a medium that perpetuates anonymity is becoming perilously difficult.

Customer interaction with financial institutions is migrating from in person, paper-transactions to remote access and transaction initiation. This migration increases the risks of doing business with unauthorized or incorrectly identified parties that could result in financial loss or reputational damage to the institution.[113]

Secure electronic service delivery is a key to providing consumers with improved, more flexible, and convenient access to financial services, and to enhancing the efficiency of banking operations. One of the challenges in implementing secure electronic service delivery is building the appropriate nonrepudiation mechanisms into the banking platform. Reliable customer authentication is imperative for financial institutions engaging in any form of electronic banking or commerce. Reliable customer authentication is imperative for financial institutions engaging in any form of electronic banking. The risks of doing business with unauthorized or incorrectly identified individuals in an e-banking environment could result in financial loss and reputational damage through fraud, corruption of data, unenforceable agreements, and the disclosure of confidential information.[114]

In a world where people increasingly do business with virtual parties they have never met and likely will never meet, authentication becomes as integral to the transaction as the exchange of goods and tender. Yet, authentication is the Achilles' heel of electronic finance. In fact, most computer intrusions are perpetuated as a result of insufficient access controls and weak authentication mechanisms. For example, in 1995, Citibank found itself in an ironic position: Its technology was not as powerful as that of a group of hackers. Citibank’s main weakness was the use of "fixed passwords" to guard its computerized cash management system. There is widespread concern, especially among those in the law enforcement community, that the financial sector is not keeping up with the security side of technological change. For example, overall industry-wide use of passwords is outdated. In fact, a 1999 General Accounting Office (GAO) report highlighted the reality of outdated access controls; it found access controls to be at the forefront of security weaknesses. Beyond the norm of gates and guards that were often inadequate, failures of logical controls––those access controls built into software—were pervasive. In the information age there are hundreds of Web sites devoted to password cracking and/or interception. The most common program used for password generation is Brute Force.[115] This widely available application generates all alphanumeric combinations until the password is deciphered.

Logical control failures create quandaries for the networks they safeguard. Logical controls are meant to require users to authenticate themselves through passwords and to limit the files and other resources an authenticated user can access and the actions that user can execute. Many of the GAO’s reviews found that managers did not identify or document access needs for individual users or groups; instead, they provided overly broad access privileges to very large groups of people.[116]

[113] Federal Deposit Insurance Corporation Financial Institution Letters. “Authentication in an Electronic Banking Environment.” Aug. 8, 2001.

[114] Federal Financial Examinations Council. “Authentication in an Electronic Banking Environment.” Aug. 8, 2001.

[115] Brute Force and other password-cracker programs are widely available: http://www.ussrback.com/passwordtools.htm.

[116] Jack Brock, Director of Government-Wide Defense Information Systems for the General Accounting Office. Testimony in 1999 before the Senate on the proposed Government Information Security Act (S1993). Brock said the GAO found that users share accounts and passwords and often post passwords in plain view, making it impossible to trace specific transactions or modifications to an individual. Unfortunately, he said, as a result of these and other access control weaknesses, auditors conducting penetration tests of agency systems are almost always successful in gaining unauthorized access that would allow intruders to read, modify, or delete data for whatever purposes they had in mind.

There are two issues here: access and authentication. Access allows those who should be able to get onto the system access for the purpose for which they are authorized. Authentication is assuring the system that the person trying to gain access or engage in a certain activity is, in fact, the person he or she claims to be and that the person is authorized to engage in the act. Used together and diligently, these processes are the most cost-effective security devices available.

Financial institutions can use a variety of access and authentication tools and methodologies to authenticate customers. Existing access control techniques and authentication methodologies involve three basic factors:

- Something the user knows (e.g., password or PIN)
- Something the user possesses (e.g., ATM card, smart card, or token)
- Something the user is (e.g., biometric characteristic, such as fingerprint or retinal pattern)

An effective access and authentication program should be implemented across the organizational structure, including affiliate entities, which requires the appropriate use of controls and authentication tools. Authentication processes should also maximize interoperability and offer consistency with the financial institution's assessment of the e-finance system risks. Before it goes online, the financial institution should examine its business processes, undergo a data classification inventory as part of its risk management analysis, and configure its rules engines and access controls to support the data classifications. This annex addresses the positive and negative attributes of passwords, tokens/smart cards, biometrics, and public key infrastructure authentication systems.

A. Access Controls: Passwords and PINs

The entry of a username or an ID and a secret string of characters such as a password or a personal identification number (PIN) is the most common and vulnerable of all single-factor authentication techniques. The effectiveness of password security depends on three characteristics: length and composition, secrecy, and system controls.

Password Length and Composition. The appropriate password length and composition depends on the value or sensitivity of the data protected by the password. Password composition standards require the use of numbers and symbols in the password sequence, as well as upper and lower case alphabetic characters. This is necessary because those attempting to compromise the system may subject systems linked to the Internet to automated attacks. These automated attacks (e.g., Brute Force) can generate millions of alphanumeric combinations to garner a user's password.

Password Secrecy. The two most frequent ways of intercepting passwords are studying the user's behavior and catching passwords in transit. First, passwords can be compromised because of the user's behavioral techniques that capture passwords as they travel across the Internet. Attackers can also exploit server and system vulnerabilities to gain access to a financial institution or its service providers to obtain password files. Given these realities, passwords and password files should be encrypted with 128-bit encryption when stored or transmitted over the Internet.[117]

Password System Controls. At a minimum, system-wide policies on computer usage and access controls should be in place. These policies should address the following: evaluation of password length and composition requirements, verification of correct logon and lockout, password expiration and revocation, encryption requirements, and activity report monitoring. Using the following security measures is essential for minimum system access controls:

[117] Refer to Annex I, under the heading Access Controls, for "Best Practices."

- Restrict the use of automatic logon features.
- Lock out users after three failed logon attempts.
- Establish an appropriate password expiration interval for sensitive internal or high-value systems.
- Terminate customer connections after a specified period of inactivity.
- Review password exception reports.
- Provide guidance on prudent, complex password selection.
- Incorporate a multifactor authentication method for sensitive internal or high-value systems.

The Inherent Weaknesses

Even with these precautions, the inherent weaknesses of passwords are technology and time. As a result of increased processor speeds, patient hackers can acquire an encrypted password file or session. A program named L0ftCrack is a random character generation program that, when used with a 1.8 gig processor, can run 1 million keyboard combinations per second.[118] The computer will execute L0ftCrack]in logical progression, thus making it a matter of time before that terminal is compromised. Note that hackers with criminal intent are patient. It may take months to set up an attack, but it takes seconds to execute a successful intrusion on a bank. Passwords can be compromised and thus provide no real level of nonrepudiation.

B. Tokens/Smart Card

A token is an authentication method that makes use of something the user possesses. Typically, a token is a two-factor authentication process, complemented by a password or a biometric as the other factor. The device itself may authenticate static passwords or biometric identifiers used to authenticate the user locally. This process avoids the transmission of shared secrets over an open network.[119]

Physical devices such as credit cards, debit cards, and telephones may contain chips that generate passwords, possess credentials, or process information when brought into contact within receivers. Tokens using the chip technology embedded in cards are known as smart cards.[120] Some are designed to hold authenticating information, and others are capable of processing information obtained in a database.

Most so called "smart cards[121]" are nothing more than a credit card sized device containing a microchip. The sophistication of the chips varies, but most all commercially available implementations are far from being secure. In considering the threat posed by criminals, it is not enough to deter the casual criminal through the inconvenience of basic security. New measures must be able to withstand the continuous and repeated efforts of a determined and well-funded adversary. Standard smart cards usually contain account numbers, encryption keys, and often additional stored information (such as biometric profiles) which can be extracted from the card and duplicated or altered. In so doing, the determined adversary can then present cloned or altered data smart cards as genuine, defeating the security and gaining access to critical infrastructure. For the use of smart cards to be effective, they must be able to dynamically respond to non-predictable challenges that defy duplication. They must not contain static, unchanging

[118] Interview with Rick Fleming, Vice President of Security Operations, Digital Defense Inc.
[119] Federal Financial Examinations Council. "Authentication in an Electronic Banking Environment." Aug. 8, 2001.
[120] Ibid.
[121] Interview with Mike Voorhees, CEO of Cryptodynamics Inc. www.cyonic.com

information such as accounts, encryption keys, biometric profiles, or other personal information. Failure in these regards leads to wide-spread distribution of system critical information that in turn becomes accessible to determined adversaries, including terrorists, and dramatically undermines the security demanded by the application in question.

A number of financial institutions use password-generating tokens to authenticate commercial customers to remotely access the institution's electronic banking system. Public key infrastructure (PKI) systems can incorporate smart cards that contain a user's credentials and digital certificate.

Therefore, the most important issues surrounding use of tokens that a financial institution needs to resolve are the following:

- Determine an appropriate expiration date and renewal and revocation process for tokens.
- Ensure that two-factor authentication processes that use tokens limit the number of login attempts.
- Educate customers about and require them to safeguard their tokens; inclu de binding agreements on the rules of use, protection, and replacement.
- Design and implement a secure process for generating and distributing tokens.

Cyonic Technology[122]

Cyonic™ technology is a core authentication system with a wide variety of applications. It is based on the use of mass-producible microchips that can be deployed in a variety of convenient consumer products. Each chip behaves uniquely in response to random or pseudo-random challenges. The system is fundamentally an authentication technology, and not an encryption technology forced to perform an authentication role. The system can function as either stand alone access control, or be used in conjunction with various biometric techniques or conventional PIN Codes. In addition to the primary authentication function, the chips can dynamically generate one-time use encryption keys and automatically encrypt a data stream to secure additional information that may be transmitted along with the authentication response.

The primary security advantage of Cyonic™ technology is that no system level information is contained on the microchip. Thus even a determined adversary, upon destructive analysis of the chip, will gain no insight into the system level functions that authenticate system users. This prevents any adversary from achieving successful account cloning, regardless of the adversary's technical or financial resources. It is superior to PKI (Public Key Infrastructure) in several important aspects. First, Cyonic™ authentication is not based on the use of certificates or stored encryption keys, which in PKI can be surreptitiously removed from the user's computer, duplicated, and presented as genuine by unauthorized personnel. Second, no key distribution or key management is required, reducing deployment and operational costs. Third, cryptanalytic techniques are not capable of revealing the behavior of Cyonic™ authentication, as Cyonic™ chips do not encrypt data for later decryption, they modify data strings with no inherent information content.

Single Factor Authentication – Cyonic™ technology allows for robust dynamic authentication with extreme security against account duplication.

Multi-Factor Authentication – For applications requiring additional protection against physical theft of, or unauthorized access to, Cyonic™ chips or devices, Cyonic™ technology can be used

[122] www.cyonic.com

in conjunction with a variety of Biometric signatures (including facial geometry, iris pattern, fingerprint, voice print, and retinal scan) as well as with conventional P.I.N. entry.

C. Biometrics

Biometric authentication techniques can grant or deny access to networks by automatically verifying the identity of people through their distinctive physical or behavioral traits. A biometric identifier represents a physical characteristic of the user. The identifier is created from sources such as the user's face or hand geometry, voice, iris (or retina), or fingerprint. Once "captured," a biometric is translated algorithmically into a complex string of numbers and stored in a database as a template.[123] Later, this template is compared to any "live" biometric presented as proof of identity. Introducing a biometric method of authentication requires physical contact with each customer to initially capture and validate the biometric information. This corresponds to the "know thy customer" mantra of the Financial Action Task Force principles.

Verification Issues

In a verification system, an individual claims an identity, typically by using a name, an ID number, or an e-mail address, and then presents biometric data such as a fingerprint to verify this claim. This biometric data is compared to the user's existing record, typically stored in a database. If the two pieces of biometric data "match" when compared, the individual's identity claim is verified. These systems are deployed when it is necessary to verify rapidly that an individual is who he or she claims to be.

Verification requires at least two things. One is that the biometric came from the actual person at the time of verification. The other is that the biometric matches the master biometric on file. If the system does not do both, it is not secure. The use of biometrics for remote login authorizations can pose problems if the data transmission is not properly encrypted. End-to-end encryption is necessary, preferably with 128-bit or greater. If the biometric is not encrypted, it is susceptible to being copied over the network as easily as any other electronic file. Finally, tuning a system properly so that "false acceptances"[124] are more common in small value transfers and "false rejections"[125] are more common in large value transfers will mitigate certain elements of risk. What is essential to the viability of biometrics as preeminent authentication technology is the test for liveliness in real time. Dorothy Denning, a tenured Professor of Computer Science at Georgetown University, reflects on biometrics:

> "What happens if someone snatches the biometric print used to validate you? Couldn't they just replay your biometric and pretend to be you? And wouldn't that make your biometric useless?" My response is, "No." A good biometrics system should not depend on secrecy. To understand why, think about how biometrics work in the physical world. Your friends and colleagues authenticate you by recognizing your face, voice, eyes, hands and so on. None of this is secret. Anyone who interacts with you sees these characteristics. Even your fingerprints can be lifted from surfaces. What makes biometrics successful is not secrecy, but rather the ability to determine "liveness." I can

[123] Security of the database is crucial. The database must be strongly encrypted and it cannot afford to be lost or inaccessible to both operational and emergency backup systems.

[124] False acceptances are how often a system would let an imposter get through.

[125] False rejections are occurrences wherein an authorized user is denied access because a system did not recognize that user.

> easily distinguish the living, flesh-and-blood you from a statue or photograph of you, or even someone wearing a costume and mask that looks like you.

Testing liveness is reasonably straightforward if the biometrics reader senses appropriate characteristics and is tightly coupled with the validation process and database of biometric prints. If the reader is remote from the validation process and database, encryption can be used to provide a secure path connecting the components. The encryption system, obviously, should protect against replays. Encryption can also be used to pass credentials from one system to another. For example, once my smart card validates a fingerprint, it may use a private signature key on the card to authenticate a user to services that use their public key for authentication. Of course, the encryption system itself requires secret keys, but in this context the secrets may be less prone to compromise because they do not have to be known by humans.

Types of Biometric Devices

Typical physical authentication methods identify you by comparing your live face against the photo, not by comparing one photo against another. For further proof, I may watch you sign your name and compare the live signature against the one on your ID card. The same principle applies in the digital world. Your biometric prints need not be kept secret, but the validation process must check for liveness of the readings. Many biometric products work this way. For instance, the iris recognition system looks for the "hippus movement"—the constant shifting and pulse that takes place in the eye. The liveness test ensures that the reading is fresh, so an adversary can not replay a previously recorded reading. This is the value of biometrics. This specter of an Orwellian central database can be avoided if users carry their own biometric iris codes on smart cards. Under this setup, if the biometric stored on the smart card matches the user’s iris scan, access to the banking network is granted. For example, there is no need for a passport database in airports; an identity can be confirmed by comparing the photograph in a passport—in this case a smart card—to the face of its bearer (iris code). This system was implemented by the Department of Defense. The Pentagon has recently begun passing out the first high-tech ID cards, which eventually will be in the hands of all 4 million military and civilian defense workers. "The credit card-size 'smart' cards will allow the holder to access secure Defense Web sites, log into computers, digitally encrypt and sign e-mail, open secure doors, get cash, buy food, and even check out weapons and military hardware. It is their passport to the electronic world," says Defense Personnel Chief David S.C. Chu. “The card has a bar code, circuit chip, and magnetic strip. It stores the person's digital picture, fingerprint, and personal identification number. It can be quickly deactivated if lost.”[126]

Recently, the city of Glendale, California, replaced its employees' password protection system with fingerprint scanners.[127]

Biometric devices are increasingly being explored as the solution to authenticate users more effectively. These imaging techniques will prevent unauthorized access. The biometric identifier represents something that the user is. The identifier can be created from sources such as the customer's or employee’s face or hand geometry, voice, iris (or retina), and fingerprints. For reasons of privacy and security, special protections must be in place when retaining, managing, and transferring biometric data. Institutions must understand the technical and legal challenges of using biometrics.

[126] www.msnbc.com. Oct. 29, 2001.
[127] Lynn Haber, www.zdnet.com. Oct. 18, 2001.

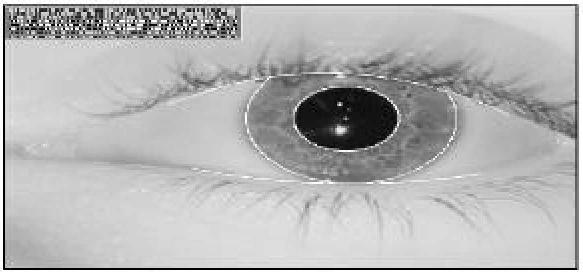
The Federal Financial Institutions Examination Council says financial institutions need to consider the following when using biometric identifiers:

- First, design systems that encrypt biometric identifiers during storage and transmission.
- Second, design and implement a secure process for capturing biometric identifiers.
- Third, limit the number of failed logons a customer can attempt.[128]

In addition to the listed considerations, biometric authentication should not be stored in a database because of the threat of compromise**.** Security of the biometric signature is paramount. Once the biological feature has been scanned, it is converted into a data string, which like any other data can be copied and presented as genuine by unauthorized personnel. Furthermore, all biometric data must be encrypted using unique one-time keys. The use of static PKI or symmetric key encryption will ultimately be compromised and must be avoided at all costs, as the same or substantially similar biometric data will be encrypted time and again. This facilitates various cryptanalytic attacks that can compromise the encryption keys. Revocation of biometrics is virtually impossible, and thus when biometric devices are employed, either they should be live-scan-activated or storage of the image should be allocated to a smart card[129]. Smart cards coupled with a biometric PIN will solve the dilemma of storage of biometric identifiers. These considerations should ensure that intuitive biometric technologies replace traditional authentication mechanisms such as passwords, PINs, and tokens.

The Iris Scan[130]

Graphic 1



$$\max_{(r,x_0,y_0)}\left|G_\sigma(r) * \frac{\partial}{\partial r}\oint_{r,x_0,y_0}\frac{I(x,y)}{2\pi r}ds\right|$$

Localizing iris boundaries by differential operators

The most effective methods of biometric authentication revolve around functional tests of liveness. Biometrics iris recognition involves identification of 266 detectable iris features that can be converted into digital code.[131] A video image of the iris of the eye is needed to produce a digitized 512-byte IrisCode™ record.[132] The image can be taken from up to 30 inches away, and therefore no physical contact is required. Iris recognition leverages unique features of the human iris to provide an unmatched identification technology. The algorithms used in iris recognition are so accurate that the entire planet could be enrolled in an iris database with only a small chance of false acceptance or false rejection. Iris recognition is based on visible (via regular and/or infrared light) qualities of the iris. Iris recognition technology converts these visible characteristics into a 512-byte IrisCode, a template stored for future verification attempts. The first step is location of

[128] Federal Financial Examinations Council. “Authentication in an Electronic Banking Environment.” Aug. 8, 2001.

[129] www.cyonic.com

[130] Iridian is the sole developer of iris-based identification technologies.

[131] Bill Rogers Editor of Biometrics Digest. Interview. Jan. 3, 2001.

[132] Ibid.

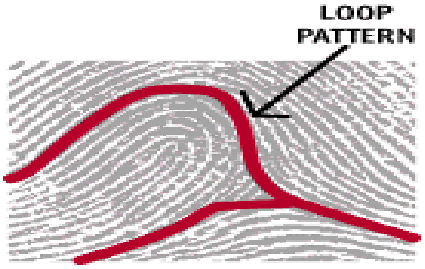
the iris by a dedicated camera no more than three feet from the eye. After the camera situates the eye, the algorithm narrows in from the right and left of the iris to locate its outer edge. This horizontal approach accounts for obstruction caused by the eyelids. It simultaneously locates the inner edge of the iris (at the pupil), excluding the lower 90 degrees because of inherent moisture and lighting issues.

No video image of the iris is retained. Instead, the eye pattern is converted into a 512-byte IrisCode record. The IrisCode is hashed and encrypted as a security measure.
No two irises are alike, even in twins and even between the left and right iris of one individual. The iris is stable over one's entire life.[133] Iris-based biometrics would be the essential safeguard for employees of financial institutions who are responsible for large-value transfers.

The Fingerprint Scan

A functional alternative to the iris scan is the fingerprint biometric. Biometric imaging electronically captures forensic-quality fingerprint images for fraud prevention, employment processing, and customer access control. These systems are crucial when it is imperative to identify people unequivocally. Fingerprint bioimaging technologies capture clear, sharp, grayscale digital fingerprints to ensure that important records are uniquely associated with one—and only one—person. The liveness test associated with these types of biometrics is that of a heat sensor under the imprint pad. The fingerprint imprint must be at least 96 degrees Fahrenheit for the device to send the authentication signal.

Graphic 2



Once a high-quality image is captured, several steps are required to convert its distinctive features into a compact template. This process, known as feature extraction, is at the core of finger-scan technology. Each of the 50 primary finger-scan vendors has a proprietary feature extraction mechanism.

More advanced optical or noncontact fingerprinting systems (known as live-scan), which normally use prints from several fingers, are the current standard for forensic use. They require 250 kb per finger for a high-quality image. Finger-scan technology also acquires the fingerprint, but it does not store the full image. It stores particular data about the fingerprint in a much smaller template, requiring from 250 to 1,000 bytes. After the data is extracted, the fingerprint is not stored. Significantly, the full fingerprint cannot be reconstructed from the finger-scan template.[134] Increasingly sophisticated mechanisms have been developed to capture the fingerprint image with sufficient detail and resolution. The technologies in use today are optical, silicon, and ultrasound.

[133] Ibid.
[134] Provided by the International Biometric Group, www.biometricgroup.com.

Optical technology is the oldest and most widely used. The finger is placed on a coated platen, usually built of hard plastic but proprietary to each company. In most devices, a charged coupled device (CCD) converts the image of the fingerprint, with dark ridges and light valleys, into a digital signal.

Silicon technology has gained considerable acceptance since its introduction in the late 1990s. Most silicon, or chip, technology is based on DC capacitance. The silicon sensor acts as one plate of a capacitor, and the finger is the other. The capacitance between the platen and the finger is converted into an 8-bit grayscale digital image. But a few firms have developed technology that employs AC capacitance and reads to the live layer of skin. Silicon-based sensors pick up electrical capacity that will serve as a functional liveness test. Sagem, Printrak, and Seimens all produce proven silicon-based fingerprint scanners.

Ultrasound technology, though considered perhaps the most accurate of the finger-scan technologies, is not yet widely used. It transmits acoustic waves and measures the distance based on the impedance of the finger, the platen, and air. Ultrasound is capable of penetrating dirt and residue on the platen and the finger, countering a main drawback to optical technology.

Voice Authentication[135]

Voice authentication technology makes use of distinctive qualities of a person's voice, some of which are behaviorally determined and others of which are physiologically determined. Voice authentication technology does not require specialized and expensive sensors or interface hardware. The entire system runs in software on conventional processors. The "infrastructure" required to implement most applications is already in place or available at low cost. In telephony applications, voice authentication technology works using the common corded, cordless, and cellular telephones. In computer-based applications, voice authentication technology works with a wide range of microphone/sound-card combinations on every standard desktop and laptop computer.

There are certain inherent drawbacks to the implementation of this technology. Because of the prevalence of high-grade recording technology, one’s voice can be recorded and used to garner a “false acceptance” from a secure network or banking system. To mitigate this risk, a “challenge response system” needs to be instituted as policy. A challenge response system asks users to verify their voice pattern by prompting them to respond with random number sequences. A randomly generated challenge response system will mitigate most of the risk inherent in recording technology. Nuance and BuyTel both provide reliable, proven voice recognition systems.

Hand Scan[136]

A hand scan reads the top and sides of the hands and fingers, using such metrics as the height of the fingers, the distance between joints, and the shape of the knuckles. Although not the most accurate physiological biometric, the hand scan has proved to be an ideal solution for low- to mid-security applications where deterrence and convenience are as much a consideration as security and accuracy. The hand scan is easy to use—the submission of the biometric is straightforward, and with proper training it can be done with few misplacements. The unit also works fairly well with dirty hands. The only problems may be with elderly clientele or people

[135] Ibid.
[136] Ibid.

with arthritic hands, who may be unable to spread their fingers easily and place their hand on the unit's surface.

Graphic 3

A hand scan is resistant to fraud. Short of casting a model of an enrolled person's hand and fingers, it would be difficult and time-consuming to submit a fake sample. Recognitions Systems Inc. (RSI), as the standard bearer of hand scan,[137] uses a template size of 9 bytes, which is extremely small—orders of magnitude smaller than in most other biometric technologies. By contrast, finger scan biometrics require 250 to 1,000 bytes, and voice scan biometrics commonly require 1,500 to 3,000 bytes. The RSI technology facilitates storage of a large number of templates in a stand-alone device, which is how many hand scan devices are designed to work. It also facilitates card-based storage, because even magstripe cards have ample room for 9-byte samples. One drawback is that the design is static; in contrast to other biometrics, which can take advantage of technological breakthroughs such as silicon development or camera quality, hand scan has remained largely unchanged for years. Its size precludes its use in most logical access scenarios, where compact design may be a prerequisite. The second drawback is cost. Hand scan readers cost approximately $1,400 to $2,000 apiece, placing them toward the high end of the authentication security price spectrum.

Facial Scan

Graphic 4

Facial scans map characteristics of a person's face into a multidimensional image. Face comparisons are made in real time. Identification involves a one-to-many comparison of an individual's face against all faces in a database in order to determine identity, and verification is characterized as a one-to-one match of an individual's face to his or her stored image for the purpose of confirming identity. Just as with finger scan and voice scan biometrics, there are various methods by which facial scan technology recognizes people. All facial scan technologies have certain commonalities,[138] such as emphasizing those sections of the face that are less susceptible to alteration, including the upper outlines of the eye sockets, the areas surrounding one's cheekbones, and the sides of the mouth. Most technologies are resistant to moderate

[137] According to the International Biometric Group, Recognition Systems Inc. is the leading manufacturer of hand scans.
[138] The International Biometric Group, www.biometricgroup.com.

changes in hairstyle, because they do not use areas of the face near the hairline. The International Biometric Group considers three firms to be in the first tier of facial scan technology: Viisage, Visionics, and BioID.

Signature Scan

Signature scan, also known as dynamic signature verification, is a biometric technology that has not seen broad use. Measuring the manner in which a user signs his or her name, password, or passphrase, signature scan examines stroke order, speed, pressure, and other factors that relate to the actual behavior of signing a tablet.

The biometrics of a handwritten signature are based not only on the shape of the signature but also on selected signature dynamics. The signature is captured, along with timing elements (speed, acceleration) and sequential stroke patterns (did the "t" get crossed from right to left and did the "i" get dotted at the very end). These unique dynamics derived from a person's muscular dexterity are usually referred to as "muscle memory." The brain, with no particular attention to detail, automatically controls these nerve impulses. These "muscle memory" dynamics yield accuracy results that are comparable to, and less intrusive than, the best alternative biometric technologies such as retinal eye scans or fingerprints. The signature and the data relevant to the transaction are collected and then bound to the signed document. There are three industry leaders in this field: Topaz Systems, Cyber-sign, and CIC.

Keystroking: BioPassword

Graphic 5

Net Nanny's BioPassword[139] 4.5 ($100 direct per seat for 50 users, $40 per seat for 4,000 users[140]) is a software alternative to hardware biometrics solutions (such as retina or fingerprint scanners), which tie a unique physical characteristic to an individual's network account to provide a positive user identification. BioPassword links the user's specific typing style and patterns to the user's password for a flexible and secure solution, without the hassle of added components. BioPassword uses two methods (or factors) to accurately identify individuals before they are granted access to critical information and resources. First, the user must know both the correct user name and password. Second, the user's typing speed and rhythm must match a biometric template. Together, these two methods dramatically reduce the chance that unauthorized users can access resources. Keystroking biometrics verify a person's typing rhythm along with his or her knowledge of a unique user ID and secret password. Other than a standard keyboard and a 300 KB 32-bit Windows DLL, no special hardware or software is needed to realize this tenfold improvement in security.

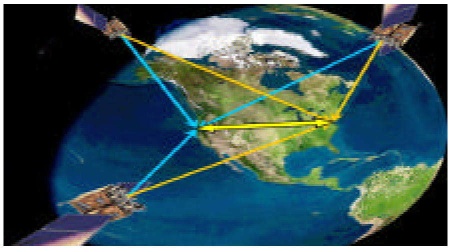
Once you are logged on, BioPassword requires you to train the system by entering a user name and password repeatedly (15 times by default). The administrator can set the amount from 1 to 20 times, but the more iterations, the better the user profile. The administrator can also establish the

[139] Net Nanny Software Inc. is the only firm to provide keystroking technology.
[140] Net Nanny Software Inc., http://www.biopassword.com/home/technology/technology_overview.asp.

accuracy required for each user with the security setting. The security setting ranges from 1 to 10, with a default of 3. The higher the number, the more accurate the user must be.[141]

Global Positioning Biometrics

Graphic 6



The Global Positioning System (GPS) is a network of satellites operated by the U.S. government that provides highly precise position and timing signals worldwide. The standard positioning service signals from the GPS are freely available for civil commercial and scientific use, and they comprise a stable and reliable standard for precise positioning and timing applications around the world. Authenticating locations involves the ability of IT personnel to authenticate the exact place and time at which access to a secure network is requested.[142] A Colorado firm, Cyberlocator Inc., is the only firm that provides location authentication services worldwide. The client's sensor acquires the signals sent from GPS satellites that are in orbit overhead at a given moment. These raw, unprocessed signals are used to create a secure location signature. This location signature, once presented to a Cyberlocator server via the network, can be used to verify that the device is located at a specific geographical point. Both the client and the server can then use this authenticated location information for any number of purposes, including network access, location-based services, and asset tracking.[143] Cyberlocator implements GPS signals for robust network access authentication and location authentication using precise position and time signatures, and for determining the precise GPS positions of wireless devices of all types without the space, weight, and power burden associated with putting a GPS processor into the wireless device.

Biometric Usage Worldwide

The following is a list of financial services providers that use biometrics to secure some of their systems: Ak Bank (Turkey), U.S. Department of Treasury, Bank United (Texas), Citibank, Dresdner Bank (Germany), Takefuji Bank (Japan), ING Direct, BACOB, Bank Belgium, Chase Bank, United Bank, Wells Fargo, NedCorp (South Africa), ABSA (South Africa), and Charles Schwab.

- Since the early 1990s, millions of South African citizens have received pension payments through a biometric smart card. Biometrics have proved very effective in overcoming problems of issuing benefits, such as a low literacy rate and the absence of a strong nationwide identification system.
- In 1998, the Philippines Social Security System commissioned Sagem to implement a large-scale ID application using biometrics. The new ID was designed to allow for the eventual electronic transmission of funds via ATMs.

[141] Jay Munro. BioPassword 4.5: Hardware-Free Biometrics. *PC Magazine,* Sept. 24, 2001.
[142] A Colorado firm, Cyberlocator Inc., is the only firm that provides these services worldwide.
[143] Input provided by Peter MacDoran, cofounder of Cyberlocater Inc.

- In July 2000, the Mexican government commissioned Visionics to create a face recognition system that would alert election officials whenever an individual attempted multiple registrations in the voting system.

Conclusion

Biometrics are the future of access controls. Biometric devices fulfill the nonrepudiation element of layered security by authenticating a user by his or her physical characteristics. Implementing biometric technologies virtually guarantees a system administrator that the person who initiated the communication or system access was who he or she should have been. The greatest obstacle that biometric technology faces lies in the acceptance of the public. Many people fear the ramifications of storing personal information in a vast database. Visions of the books *1984* and *Gattaca* spring to mind as those who fear centralized governance reject these methods of authentication. The e-financial world must evolve past our fears of "big brother" in order to face the security challenges that will face all "virtual" industries in the years to come. Authentication is the gargantuan cyber-loophole that is exploited more often than not in order to gain access to others' computer systems. Those who advocate the use of PKI should recognize that biometric technologies are superior to PKI in terms of the security levels they are able to provide during the transfer of data. "Whereas PKI's security rests entirely on the management of the private key, biometrics register users by their inherent characteristics, thus making it more difficult to assume the identity of the user."[144] PKI encrypts data from end to end, which does not solve the authentication issue. For instance, if a user stores the information to access the private key of PKI on a PIN, this can be compromised with relative ease, particularly if the PIN is stored on a computer's hard drive or a personal digital assistant (PDA).

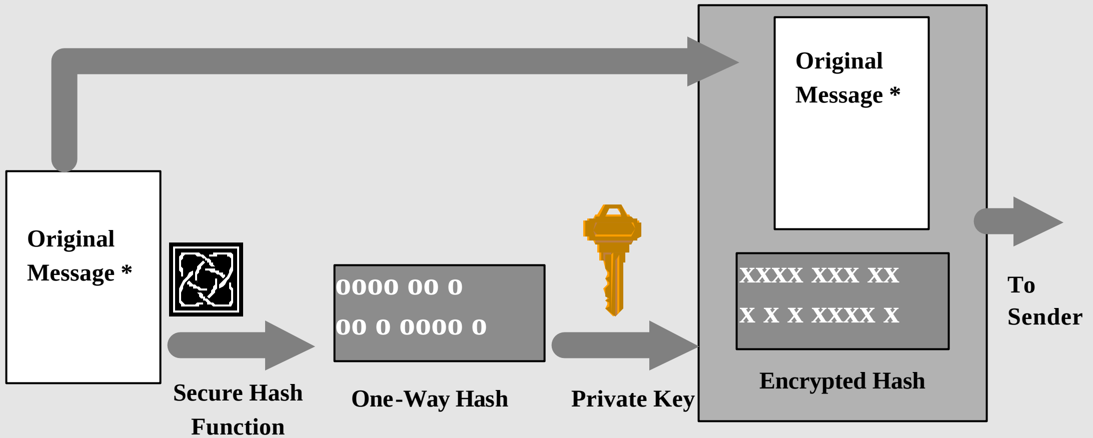
D. Public Key Infrastructure

In a PKI system, the fundamental function of a certification authority is to verify the association between a particular person and a particular public/private key pair. A certification authority functions in effect as an online notary, a trusted third party that confirms the identities of parties sending and receiving electronic payments or other communications. Because banks already have a traditional role as a trusted third party in financial and commercial transactions, they are in this respect a natural fit for the certification authority business.[145] PKI allows users to interact with other users and applications, obtain and verify identities and keys, and register with certificate authorities (CAs), which act as trusted third parties and vouch for users and their identities. This technology has existed for more than 20 years. The certificates issued by the CAs are an encrypted package containing the identity of the end recipient together with details of the chain of trust through which the recipient was recognized. It is important to note that these certificates require compliance to policies and procedures where a high level of assurance can be provided as to the authenticity of the identity of the end recipient and that the certificate is available only for his or her use. The international standard for the implementation of functional PKI networks was a creation of the International Telecommunications Union.

[144] Interview with Raj Nanavati, partner at International Biometrics Group, Dec. 3, 2001.
[145] Office of the Comptroller of the Currency (OCC) Bulletin 99-20 on certificate authority systems.

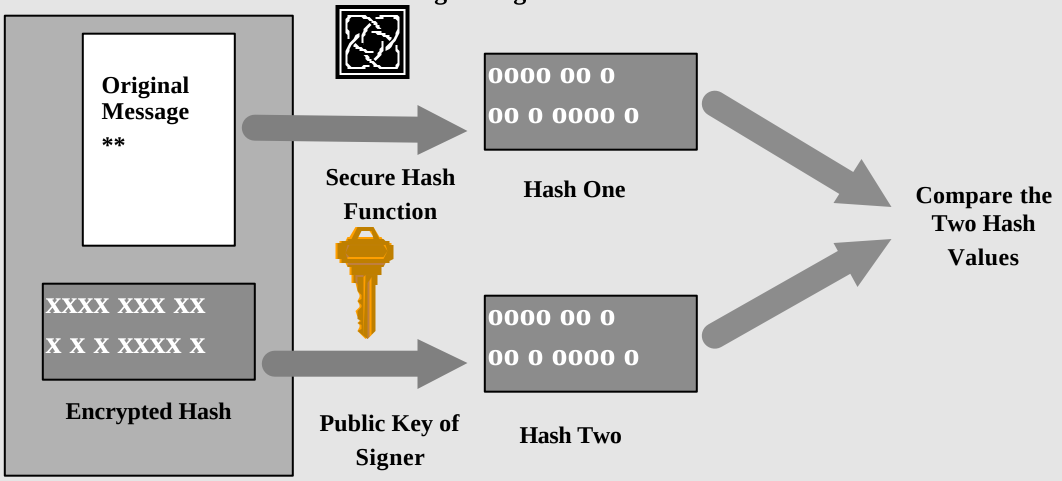
Box 1. Digital Signatures, Certificates, Certificate Authorities, and Public Key Infrastructures

Illustration A - Generation of Digital Signatures



The sender of the message would encrypt this hash using his or her *private* key and send it together with the original message to the recipient. In order to verify that the message did indeed come from the right person, the recipient would perform the following: (i) use the sender's *public* key to decrypt the hash; (ii) generate a separate hash using the original message; and (iii) compare the two hashes.

Illustration B - Digital Signatures Verification



In this manner, digital signatures also enable nonrepudiation of transactions, because the recipient is able to prove that the message (in its entirety) could only have originated from the sender, because he or she is the only person who can hold the private key. It should also be noted that this process is normally performed automatically by the programs that support digital signatures.

Source: Hong Kong Monetary Authority.

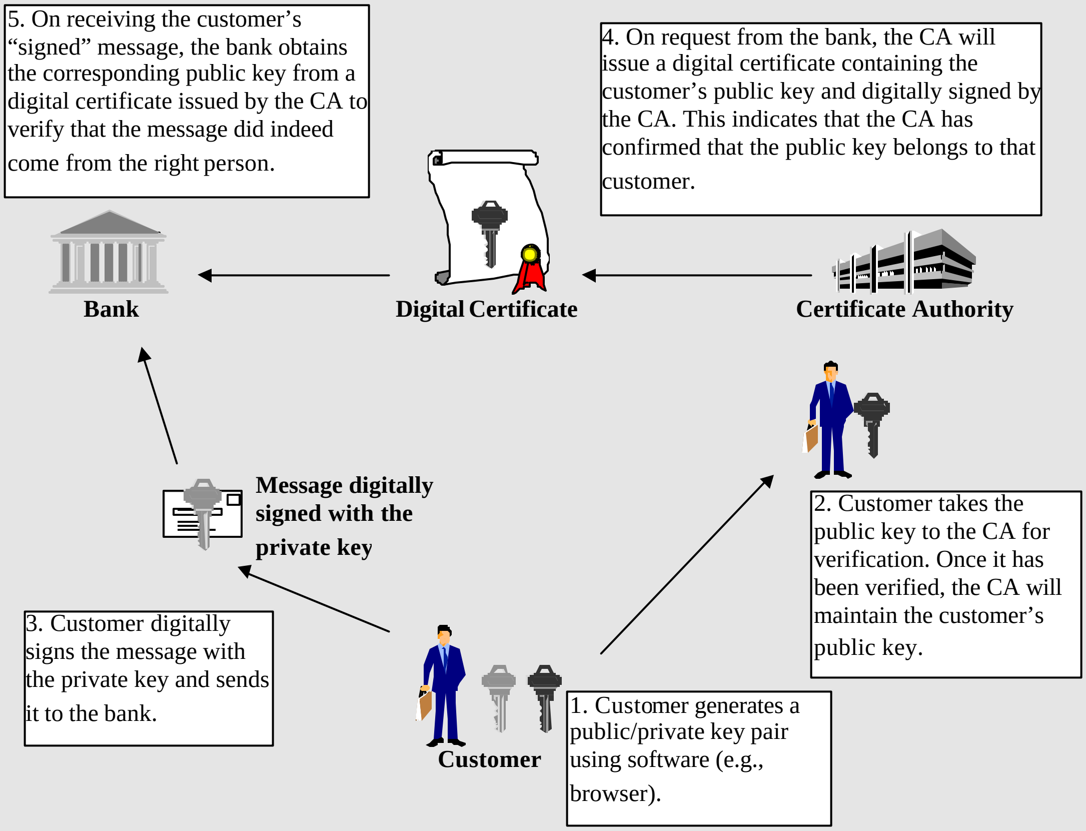
The most widely accepted format for certificates is defined by the ITU-T X.509 international standard. X.509 is viewed throughout the information technology industry as the definitive reference for designing applications related to public key infrastructures (PKIs). The elements defined within X.509 are widely used—from securing the connection between a browser and a server on the Web to providing digital signatures that enable one to conduct electronic transactions with the same confidence one has in a traditional paper-based system.[146] This new

[146] International Telecommunications Union, http://www.itu.int/ITU-T/news/sg7-x509.html.

edition, developed in close cooperation with ISO/IECand the ISOC/Internet Engineering Task Force, supersedes and replaces the 1997 publication.[147]

- It contains specific enhancements to public key certificates to support the correct processing of certification paths that involve multiple certification authorities within multiple enterprises, as well as enhancements in the area of certificate revocation.[148]
- It contains a significant enhancement to attribute certificates and definition of the framework for privilege management infrastructure (PMI). Attribute certificates will play a major role in globally addressing the complex security issues of access control and authorization. They are a standardized mechanism for defining user access privileges in a multivendor and multiapplication environment. These issues are just now coming to the attention of IT planners, as organizations move their mission-critical business relationships to the Web.

Box 2. Authentication Using Digital Certificates and Certificate Authorities



1. Customer generates a public/private key pair using software (e.g., browser).

2. Customer takes the public key to the CA for verification. Once it has been verified, the CA will maintain the customer's public key.

3. Customer digitally signs the message with the private key and sends it to the bank.

4. On request from the bank, the CA will issue a digital certificate containing the customer's public key and digitally signed by the CA. This indicates that the CA has confirmed that the public key belongs to that customer.

5. On receiving the customer's "signed" message, the bank obtains the corresponding public key from a digital certificate issued by the CA to verify that the message did indeed come from the right person.

Public key infrastructures: For an Internet banking transaction, it is possible for a bank to act as the CA for its own customers. Banks already perform a certain amount of verification on their customers before an application is approved.

Source: Hong Kong Monetary Authority.

Digital signatures use asymmetric-crypto systems in ensuring that they efficiently discharge the role of a good signature system. Both a private key and a public key are involved. The private key

[147] Ibid.
[148] Ibid.

is used for creating a digital signature or transforming data into a seemingly unintelligible form. The public key is used for verifying a digital signature or returning the message to its original form. The authentication process consists primarily of key exchange. A key is a numerical value that, according to a protocol (X.509) used by both parties, allows the receiving party to decrypt the encrypted messages. After a successful key exchange, the system sets up a security association, which permits secure (encrypted) communication between the two parties. The private key is known only to the signer and is used to create the digital signature. The public key is more widely known and is used by a relying party to verify the digital signature. If many people need to verify the signer's signature, the public key must be available or distributed to all of them, perhaps by publication in an online repository or directory where it is easily accessible. The private key and the public key are mathematically related. However, in an efficient and secure system, one cannot derive the private key from knowledge of the public key. Thus, though many people may know the public key of a given signer, they cannot find out the signer's private key and use it to forge digital signatures.

Those who advocate the use of PKI should recognize that biometric technologies are superior to PKI in terms of the security levels they are able to provide during the transfer of data. "Whereas PKI's security rests entirely on the management of the private key, biometrics register users by their inherent characteristics, thus making it more difficult to assume the identity of the user."[149] PKI encrypts data from end to end, which does not solve the authentication issue. For instance, if a user stores the information to access the private key of PKI on a PIN, this can be compromised with relative ease, particularly if the PIN is stored on a computer's hard drive or a personal digital assistant (PDA).

The effectiveness of PKI, however, depends on the integrity of the public key and on the fact that some verification has initially been performed to associate the public key to its rightful owner. This can be achieved with the use of *digital certificates*.[150] A digital certificate is an electronic document that binds an identity to a public key. It contains certain information, such as the name of the owner, the validity period of the certificate, and the public key. This set of information is verified by a certificate authority, which digitally signs the certificate using the CA's private key to affirm the integrity of the certificate. A person who obtains a public key from a certificate issued by a CA can rely on the fact that the CA has performed the necessary verification of the identity of the key owner and rely on this knowledge to transact using the keys. The CA thus acts as a "trusted" party and itself must maintain a very high level of security to protect its own private keys and to maintain the list of valid certificates issued. It is possible for a bank to act as the CA for its customers.

Banks already perform a certain amount of verification on their customers before an application is approved. But a bank acting as a CA would need to establish strong security systems, because it takes on the responsibility for maintaining the public keys of its customers. Such a scheme might be considered adequate where customers deal only with their own bank. But this could become impractical for services across banks if individual institutions adopt different standards with no interoperability between them. Customers might also be required to hold a number of certificates associated with different banks. These factors can detract from the convenience that electronic banking aims to offer. Even outside the area of banking, this is an issue at the broader level of electronic commerce, where public key systems would also be the preferred means of authentication. Root certificate authorities form the very foundation on which public key infrastructures are built. Successful vendor interoperability testing is an important milestone in a

[149] Interview with Raj Nanavati, partner at International Biometrics Group, Dec. 3, 2001.
[150] Hong Kong Monetary Authority (HKMA).

PKI that purports to be open. Trust is the key element in any certification system. Because CAs may not be trusted by everyone who uses a given PKI, the CAs themselves are certified by other CAs. Moreover, because these “superordinate” CAs may have a different scope or set of requirements, they may not be trusted by everyone in a given PKI either. At some point, then, each PKI has a single trusted "root," from which all certification disseminates. Examples of root CAs are the VeriSign Digital ID Center and the U.S. Postal Service’s Root CA. The aim of a public key infrastructure is to allow the adoption of common standards by all certificate authorities. The process of cross-certification can be further simplified by having a single entity take on this role. Such a “root authority” would establish the standards and certify subordinate CAs that comply with its standards. A root authority would also be in a position to establish cross-certification with other recognized root authorities overseas.

The Role of PKI in Singapore and Hong Kong

Singapore. The Electronic Transaction Act (ETA) of 1998 has put in place a voluntary licensing scheme for certification authorities. It is possible to verify the identity of a person online through digital certificate and signature, if approved by the certificate authority that certifies a given public key with a given individual. Netrust is the first CA to issue keys for digital signatures in Singapore. PKI applications exist in Singapore. The ETA facilitates the setting up of authentication and certificate authorities. It also sets the guidelines that will be observed by all CAs. The liability of a licensed CA is limited under the ETA.[151] “The CA will not be liable for any loss caused by reliance on a false or forged digital signature of a subscriber so long as the CA has complied with the requirements under the act and the regulations.”[152] There are regulations stipulating when a digital signature will qualify as a secure digital signature (i.e., legally binding).

Hong Kong. The Hong Kong Postal Authority has established a PKI and, since January 31, 2000, has acted as a public certification authority in Hong Kong. The Electronic Transactions Ordinance (ETO) gives the same legal recognition to digital signature that has been accorded to its paper-based counterpart. Different types of electronic signatures can be used in commercial dealings, based on agreement between the contracting parties. Where a digital signature is chosen to meet a legal requirement for a signature, only a digital signature (electronic signature based on the PKI technology) supported by a recognized certificate can be used to satisfy the requirement.

- The ETO establishes a voluntary recognition scheme for CA operations in Hong Kong. Government recognition is given only to those CAs that have attained a specified level of security and trustworthiness.
- A Certification Authority Recognition Office has been established to process applications for CA recognition and to monitor compliance by recognized CAs with the provisions of the ETO and the code of practice issued by the government.
- The electronic service delivery scheme launched by the government in December 2000, whereby public services involving financial and nonfinancial transactions are performed via the Internet, also provides for the use of PKI and digital certificates for securing transactions between the public and the government. The establishment of a public CA through the Hong Kong Postal Authority provides the foundation for the public key infrastructure.

In fact, since the first Digital 21 IT Strategy back in 1998, the HKSAR Government has made substantial progress to the improvement of Hong Kong's information security infrastructure, e.g.the enactment of Electronic Transactions Ordinance, development and promotion of PKI and

[151] New Singapore PKI Regulations Online, http://www.fitug.de/debate/9902/msg00182.html.
[152] Ibid.

e-cert, increase in the number of recognized certification authorities (total of 4 CAs in Hong Kong now),establishment of the Hong Kong Computer Emergency Response Team Coordination Centre(HKCERT/CC) and the Inter-departmental Task Force on Computer Related Crime.

Best Practices for Administration of PKI

PKI builds on, but is distinct from, public key cryptography. It enables an encrypted communication to be wrapped in an electronic "envelope," so that the recipient can verify its origin and be sure that it has not been altered while en route. The basic encryption of the communication gives confidentiality. The future of PKI lies in scaling down expectations and focusing on closed domains, such as large enterprises and groups of trading partners. PKI is one step in the layered process of functional and effective layered security.

The PKI must be structured to be consistent with the types of individuals who must administer the infrastructure. Providing these administrators with unbounded choices not only complicates the software required but also increases the chances that a subtle mistake by an administrator or software developer will result in broader compromise. Similarly, restricting administrators with cumbersome mechanisms will cause them not to use the PKI. Management protocols are *required* to support online interactions between PKI components. For example, a management protocol might be used between a CA and a client system with which a key pair is associated, or between two CAs that issue cross-certificates for each other.

- Before specifying particular message formats and procedures, one must first define the entities involved in PKI management and their interactions (in terms of the PKI management functions required). Then these functions should be grouped in order to accommodate different identifiable types of end entities. The entities involved in PKI management include the end entity (the entity to whom the certificate is issued) and the certification authority (the entity that issues the certificate). A registration authority *may* also be involved in PKI management.
- In general, the term "end entity," rather than subject, is preferred in order to avoid confusion with the field name. It is important to note that the end entities here will include not only human users of applications but also applications themselves (e.g., for IP security). This factor influences the protocols the PKI management operations use; for example, application software is far more likely than human users to know exactly which certificate extensions are required. PKI management entities are also end entities in the sense that they are sometimes named in the subject or Subject/Alt/Name field of a certificate or cross-certificate. All end entities require secure local access to some information—at a minimum, their own name and private key, the name of a CA that is directly trusted by this entity, and that CA's public key. The form of storage will also vary from files to tamper-resistant cryptographic tokens.
- Any CA system used for banking transactions should be a "closed system restricted to participants that have agreed to meet the minimum operating standards, to operate according to common business practices, and to abide by that provider's rules and regulations."[153]

[153] Excerpt taken from the Office of the Comptroller of the Currency's November 1999 Statement of Conditional Approval #339, www.occ.treas.gov.

Validation of Certificates[154]

When relying on certificates, end-users must be able to determine whether the certificate is valid. There are a number of ways to accomplish this.

Certificate revocation lists (CRLs). Most early PKI models relied on CRLs to determine certificate validity. If a certificate was no longer valid, the CA would revoke it and add it to the CRL. CRLs are somewhat static—if a certificate has been revoked since the last CRL was issued, there is no record of the fact, nor will there be until the next CRL is issued. Quick and permanent revocation of certifications is the most important attribute to truly functional CAs.

CRL distribution points. In some cases, a CA may want to partition its CRLs to identify the reason a certificate was revoked. The Open CRL Distribution Process (OpenCDP) was designed to support this concept.

Online Certificate Status Protocol (OCSP). Just as credit booklets have given way to online verification scanners, CRLs have begun to give way to real-time verification mechanisms such as OCSP. OCSP relies on the existence of real-time "responders," which can be queried to check a certificate's validity. OCSP has also given rise to similar protocols, such as Real-time Certificate Status Protocol (RCSP), which attempt to improve on OCSP.

The Inherent Technical Weaknesses of PKI

Two important security questions arise about PKI. First, where are the certificates to be held? PC browsers are not safe as storage points. Web sites like http://astalavista.box.sk instruct their patrons on how to gain access to computers and networks by manipulating cracks in their browsers' programming codes. These problems make proliferation of smart cards necessary. Second, should the existing certification system be revamped to include the ability to track/audit all those who use it? In sum, operation over a wide area network forces individuals to accept certain losses to their privacy in order to ensure the security of their virtual identity and that of their business. Creating audit logs is critical in maintaining the system. Audit logs are tools that can help monitor the certification process by tracking what a certified user does; the user can and will sign the log for an audit trail that cannot be forged. "It is this feature that is the basis of the legal enforceability of the electronic contract," says Verisign.[155] Implementing biometric technologies would lessen the chances of an unauthorized user manipulating the passwords of an unsuspecting individual. The audit log's accuracy would be much less fallible if biometric technologies were implemented as private keys.

The largest drawbacks associated with PKI deployment came from imprecise definitions of trust. Some Certificate Authorities (CA) defined trust to mean that they maintain security only when handling their own private keys; it had no bearing on their procedures for the handling of other companies' keys. No authority has the power to grant accreditations, leaving the risk in the hands of the verifier of the certificate. The inherent problem here stems from the worldwide saturation of certificate authorities who maintain the power to handle private keys. Another issue is protection of the private signing key. Most enterprises will not own a secure computing system with physical access controls and air wall network security. The key is potentially open to attack by viruses and other malicious programs, and it could be misused while vulnerable, with disastrous consequences.

[154] Contributed by Verisign Inc.
[155] Ibid.

Existing Vulnerabilities Associated with PKI

A nefarious individual can manipulate PKI systems by executing the following attack methodologies:

1. One can place a Trojan virus in a word document and send it to a user's private key file.
2. One can flood PKI with fictitious user IDs and names. The net effect is that any typo gets the user one of your prepositioned keys. One then decrypts their messages and forwards them encrypted to the intended recipient.
3. One can use social engineering to get a person to encrypt items that you provide them, and use it to get their private key.
4. One can break into a server that holds public keys and change them to ones that you specify.
5. One can crash a few key servers that form the base of a PKI tree for the users you want to defeat and they will only be able to communicate in plain text.
6. One can generate tons of key traffic so that the system overloads with requests and shuts down.
7. One can use the "Van-Eck attack" to observe secret messages after they are encrypted.
8. One can use video-viewing to observe the keyboard of users as they type in their keys or messages before encryption.
9. One can use a parallel processor to break keys of limited length. This has been successful against systems of common key lengths.
10. PKI in its purest sense is not an authentication technology but rather an encryption technology. All private keys are protected by either a token or a password, both of which are easily compromised.

Annex III. E-Security in the Case of Wireless[156]

Wireless networks are available in three basic formats: high-powered microwave systems used by telephone companies for long-haul, line-of-sight communications; CDMA/TDMA/GSM (Groupe Spécial Mobile), cellular and PCS[networks used for wireless phones and personal digital assistants (PDAs); and wireless LANs using the 802.11b protocol. While all of these are common throughout the world, they all suffer from the same basic security flaw. They use radio frequency (RF) technology to transmit their information. The result can be compromising of their transmissions.

Wireless networks (WLANs) have seen explosive growth in their deployment. With cost savings at an all-time high and with the simplicity of installation, WLANs have been deployed rapidly. Wireless networks were supposed to do what traditional Ethernet LANs do without cables. Convenience for the customer is paramount in the proliferation of wireless. Wireless technology is built around the 802.11b IEEE standard in the United States and the GSM standard in Europe. This annex discusses the security issues raised by both of these forms of cellular technology and provides a glimpse into the future of third generation wireless (3G).

Table 1. The Wide Range of Mobile Services[157]

Cellular and PCS Services	GSM (TDMA), CDMA, digital cellular, and 3G
WLANs	The 802.11a and b standards, as well as the new 5.7 GHz wireless LAN band.[158]
Satellite Systems [159]	Ka-band desktop services[160] and the Astra satellite network[161]

A. 802.11b Vulnerabilities: An American Phenomenon

Wireless LANs make use of the IEEE 802.11b technology, which is a system that transmits and receives in the 2.4GHz range and is capable of a maximum network capacity of 11Mbps. WLANs implement the Wireless Equivalent Protocol (WEP), which was designed to offer the same security features as a physical wire: confidentiality, access control, and data integrity. 2001's Black Hat Briefing made public that hackers have a multitude of ways in which they can crack, interject, or modify WEP messages on a wireless network. There is a particular problem with devices using the 802.11 wireless network standard. The encryption can easily be broken, and once broken it can provide easy access to corporate networks for anyone listening in.

[156] This annex is an excerpt from *Mobile Risk Management* written by Tom Kellermann**,** Data Risk Management Specialist, World Bank, April 2002.
[157] Provided by Dr. Joseph Pelton, Executive Director of the Clarke Institute.
[158] This new LAN is now being used for many new applications that involve financial transactions, from toll highways in Europe to banking and bank-to-bank transactions.
[159] These systems provide for both trunk-line Internet transmissions and digital video broadcast services for video streaming and cache updating, as well as direct access services.
[160] These systems began in Europe.
[161] Hughes Spaceway System.

Furthermore, if a wireless gateway is located on the corporate Ethernet network, that network will broadcast all the data passing through it over the airwaves. If someone cracks the encryption, that person can intercept everything. But the immediate points of vulnerability are the mobile devices themselves, including notebooks, which tend to be poorly protected and which often contain sensitive but unencrypted data. The danger to financial and corporate networks is very real. The 802.11b standard includes a provision for WEP encryption. Depending on the manufacturer and the model of the NIC card and access point, there are two levels of WEP commonly available— one based on a 40-bit encryption key (also called 64-bit encryption because it uses a 24-bit initialization vector (IV) and is considered very insecure) and the other using a 104-bit key plus the 24-bit IV (also called 128-bit encryption).

A recent technical paper titled "Weakness in the Key Scheduling Algorithm of RC4"[162] laid out several fundamental flaws in how the RC4 encryption algorithm was used in the WEP encryption scheme. This paper proposed a method for determining the master WEP key that would allow a hacker to pose as a legitimate user of the network. Shortly after that, a program called AirSNORT appeared on the Internet.[163] AirSNORT takes advantage of the flaws outlined in the "Weakness" paper and can, after monitoring a wireless network for some time, discover the WEP key. The essential problem with WEP is that the underlying clear text message used to frame the information in the 802.11 header is predictable and repeatable. Given enough cipher text coupled with clear text, a cryptographer can find the key.

When designing a wireless network, one should keep in mind a number of important security concerns. These are the seven basic categories of wireless network security risks:[164]

1. **Insertion attacks** – The intruder attempts to insert traffic into your network, typically through an unsecured mobile access point.
2. **Jamming** – This is a DoS (denial of service) attack, where the attacker tries to flood the radio frequency spectrum of your wireless network by broadcasting packets at the same frequency as your network.
3. **Encryption attacks** – The IEEE 802.11b wireless network standard uses a WEP encryption method. This standard uses weak encryption and initialization vectors and has been cracked successfully many times.
4. **Traffic interception and monitoring** (war driving) –Wireless packets using the 802.11b standard have an approximate transmission distance of 300 feet. This means that anyone with the proper standard equipment can receive that signal if he or she is in transmission range. Equipment to extend that range further is easily available, so the area of interception can be quite large and hard to secure properly.
5. **Mobile node to mobile node** – Most mobile nodes (laptops, PDAs) are able to communicate directly with each other if file-sharing or other TCP/IP services are running. This means that any mobile node can transfer a malicious file or program rapidly throughout your network.
6. **Configuration issues** – Any wireless device, service, or application that is not correctly configured before installation and use can leave an entire network at risk. Most wireless devices and applications are preconfigured to accept any request for services or access. This means any passing mobile client can request and receive telnet sessions or ftp.

[162] Scott Fluhrer, Istak Mantin, and Adi Shamir.

[163] Input provided by Rick Fleming, Vice President of Security Operations, Digital Defense Inc.

[164] Chris Bateman of CERT Analysis Center contributed the seven wireless vulnerabilities.

7. **Brute Force attacks** – Most wireless access points use a shared password or key for all devices on that network. This makes wireless access points vulnerable to Brute Force dictionary attacks against passwords.

War Driving

Industrial espionage and white-collar crime have reached new heights with the advance of new technologies. War dialing, the hacking practice of phoning up every extension of a corporate phone network until the number associated with the firm's modem bank is hit upon, has been replaced by *war driving.* War driving involves motoring between targeted financial institutions and corporate headquarters with a laptop fitted with a WLAN card and trying to record network traffic (sniffing). According to Dave Thomas, the Chief Investigator of the FBI Computer Crimes Division, war driving is a widespread phenomenon that jeopardizes the security of all institutions and corporations that implement WLANs.

When testing and deploying WLANs, a network administrator may find that the institution's laptops can only connect to the access points within a certain distance and may therefore assume that the signals do not travel beyond this point. This is a flawed assumption. In fact, these signals may travel for several thousand meters if there is nothing in the way to deflect or interrupt the signal. The reason for this misconception is that the small antennas in the laptops cannot detect the weaker signals. But if external antennas are used, the range can be vastly extended. The wireless segment is usually omnidirectional, so a potential adversary need not gain physical access to the segment to sniff (or record) the packet traffic. As a result, WLANs are susceptible to message interception, alteration, and jamming.

These considerations raise the issue of how to secure wireless networks better. This will be as critical as securing fixed-line Internet systems in the emerging markets, as highlighted above. Each of these security breaches and associated risks can be minimized or negated with the proper use of security policy and practices, network design, and system security applications, and with the correct configuration of security controls.

15 Steps to Securing WLANs

Wireless network security is much like the physical security at the entrance of a building. Someone with enough interest, resources, and time is going to be able to gain access. First and foremost, it is important to treat your wireless network as though it were a publicly accessible network. A system administrator should not make any assumptions that traffic on that network is private and secure.

The following security recommendations, compiled from a host of industry leaders, offer some simple rules that can provide a foundation for securing a WLAN:

1. **Create an institution-wide policy on wireless devices**. Tailor the corporate security policy to address network use guidelines.
2. **Track how many employees have WLANs at home**. These remote access users need to be monitored to eliminate unauthorized wireless access points.
3. **Disable all unneeded services and applications on each client and server**. Typically, all services and applications that are not known or in use *should be disabled.*
4. **Change the default settings of your product**. Many administrators make the mistake of not changing any of the server set ID (SSID) or IP address information for their access

points. Do not change the SSID to reflect your company's name, divisions, or products. Because this information is broadcast by the access point, once hackers have broken WEP, they know exactly whose network they are accessing.

5. **Change the default password on your access point or wireless router**. Hackers often know the manufacturers' default passwords and will try them first.
6. **Plan your coverage to radiate out to the windows but not beyond**. As you do your site survey for access point deployment, think about locating the access points toward the center of your building rather than near the windows. If the access points are located near the windows, a stronger signa l will be radiated outside your building, making it easier for people to find you.
7. **Provide directional antennas for wireless devices**. Most wireless devices use omnidirectional antennas; these antennas allow for systematic sniffing (recording) of all communications. Directional antennas, coupled with a 2.4-gig or higher frequency, will lessen the propagation of the signal.
8. **Turn WEP on** and manage your WEP key by changing the default key and changing the WEP key every week.[165]
9. **Use VPN tunneling between the network firewall and the wireless**. Though it would require a VPN server, the VPN client is already included in many operating systems, such as Windows 98 Second Edition, Windows 2000, and Windows XP.
10. **Deploy a network-based intrusion detection system (NIDS) on the wireless network**.[166]
11. **Deploy enterprise-wide antivirus software on all wireless clients.**
12. **Employ two-factor authentication.** The two-factor approach mitigates a tremendous amount of risk. Two-factor authentication is best employed in one of two ways. One is through token-based smart cards that store a biometric record.[167] The other is through the use of radius servers, which authenticate the machine to the network. A radius server permits association with your access points. A user connects to the radius server merely for authentication to the other servers. One can implement a biometric to initialize the server, thus abiding by the two-factor authentication mantra. Radius servers[168] act as a guard would in a lobby, authorizing passage to the rest of the building.
13. **Consider using a wireless firewall gateway**.[169] This device operates as a standard dual-homed firewall with the wireless network on one side and the trusted network on the other. The firewall has security software such as IPSEC or some other VPN enabled-software, and only after authenticating to that software can access be granted to the internal network. The firewall rules may also be used to limit where traffic originating from wireless networks may traverse. Make sure the network firewall is between all wireless access points and the internal network or Internet.

[165] Input provided by the NIPC, http://www.nipc.gov/publications/nipcpub/bestpract.html.

[166] Input provided by Chris Bateman of CERT Analysis Center.

[167] Bateman recommends the e-thenticator, which is a thumbprint biometric scanner that stores the image on a smart card. The cost is roughly $100 per unit.

[168] RADIUS, or Remote Authentication Dial-In User Service, is an authentication service that verifies user information and, once verified, allows users to access certain network services. Part of what RADIUS can provide is encrypted communication between the remote client and the RADIUS server. Virtual private networks (VPNs) work in a similar manner but tend to operate on a network-to-network connection instead of the remote host-to-network method of RADIUS. Once the remote computer is authenticated and connected to the internal network via a RADIUS server, it operates as if it were physically located near and connected to the network. In other words, the encryption provided by the RADIUS server is only between the RADIUS server and the client machine, not over the network as a whole. Rick Fleming, Vice President of Security Operations at Digital Defense Inc., has stated that “Cisco’s Aeronet Tacacs Server is premier for this service.”

[169] Rick Fleming, Vice President of Security Operations, Digital Defense Inc.

14. **Disable DHCP [spell out?]and use static IP addresses for your wireless NICs**.[170] Also change the default IP address range for your wireless network from the manufacturer's default.
15. **Purchase access points that have "flashable" firmware only**. A number of security enhancements are being developed, and you want to be sure that you can upgrade your access point.

Proper System Administration and Auditing

The proper administration of a wireless network is one of the main components of achieving reliable security. The system administrator should routinely perform the following tasks:

1. Reconfigure any wireless device from its factory settings before it is deployed on the network. Turn off all unnecessary services.
2. Obtain the latest security fixes from the vendor and install appropriately before deployment.
3. Review all firewall logs weekly, and scan critical host logs daily.
4. Review all ACLs and user accounts on a monthly basis. System administrators should remove all access privileges for terminated employees.
5. Scan automatically all downloads, using enterprise antivirus software.
6. Set password content and length policy to at least 10 alphanumeric characters.
7. Review all IDS logs weekly.
8. Maintain an inventory of all mobile devices.
9. Prohibit all unauthorized wireless devices.
10. Develop a standard wireless access point configuration, and use it on all nodes.[171]
11. Change all access point SSIDs (server set IDs). These are the shared passwords that come factory-installed on all wireless access points.[172]
12. Disable SNMP community passwords on all access points.
13. Enable 128-bit WEP encryption.
14. Move or encrypt the SSID password and the WEP key. Most wireless clients store the SSID password and a shared WEP key in the Windows registry file.

Also, at least biannually, a penetration test or risk assessment should be performed. The results should be used to drive new policy, equipment, and network configuration changes. A wireless network is relatively easy to test for vulnerabilities, and a war-driving system can identify key security risks cheaply and quickly. A more comprehensive risk assessment methodology called OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) is available from the Software Engineering Institute at http://www.cert.org/octave. There are several freeware war-driving software packages available now, such as AirSNORT), and several made by wireless technology companies (e.g., IBM's WSA) that will provide the organization with a solid picture of the network's vulnerabilities.[173]

B. The European Cellular Standard: GSM

In 1982, the Conference of European Posts and Telegraphs (CEPT) formed a study group called the Groupe Spécial Mobile (GSM) to study and develop a pan-European public land mobile

[170] Ibid.
[171] Contributed by Chris Bateman of CERT.
[172] Ibid.
[173] Ibid.

system. Today, GSM is the world's most widely deployed and fastest growing digital cellular standard. GSM subscribers worldwide number nearly 600 million, more than two-thirds of the world's digital mobile population. The numbers are increasing by four new users per second. GSM covers every continent, being the technology of choice for 400 operators in more than 170 countries.[174] But this is only the beginning of the wireless revolution. The industry predicts more than 1.4 billion GSM customers by the end of 2005.[175]

GSM phones have a small smart card inside them that holds the identity of the cell phone. This small smart card is called a Subscriber Identification Module (SIM). The SIM must keep the identity inside secret and uses cryptography to protect it.

The North American GSM system operates at 1900mhz in conjunction with digital PCS services. The data services associated with GSM are Short Message Service (SMS), Analog Cellular Switched Data (CSD), and General Packet Radio Service (GPRS).[176] Most European cellular carriers use a form of GSM, in either 900mhz or 1800mhz.[177] Europeans also have the option of using High Speed Circuit Switched Data (HSCSD), which combines several channels into a single channel capable of 38.4 KBPS. GPRS is also available in most countries.[178]

GSM Vulnerabilities

SIM Card Vulnerability. In both European and American GSM systems, the network access method is the same. Removable smart cards in the phone (SIM cards) are used to store phone numbers, account information, and additional software such as wireless Web browsers. The data on the cards is encrypted, but the COMP128 algorithm that protects the information on the card has been compromised, making these cards susceptible to duplication. War driving is not a substantial issue for cellular subscribers using GSM. Regardless of frequency, cellular signals can easily be jammed. There is a widely known method for recovering the key for an encrypted GSM conversation in less than a second using a PC with 128 MB of RAM and 73 GB of hard drive space.

The security of GSM phone technology is limited. It is possible to clone GSM SIM cards. The hack attack is possible because critical algorithms are flawed, making it possible to dump the contents of the SIM cards and then emulate them using a PC.[179] This latest problem could render GSM phone conversations totally insecure. For a bank, there are other issues. For example, a remote teller machine could be tricked into communicating with a fake mobile tower because it cannot reach a real one. This would allow the perpetrator to remotely control the transmissions of funds via the teller machine. In the figure below, a modified GSM cell phone and laptop are being made to act as a base station. All that is necessary is to make a few software and hardware modifications to the phone and to be within closer range than the actual tower. The mobile phone must authenticate itself to the base station, but the station does not have to authenticate to the phone at all. To eliminate the unknown variables, a fixed RAND challenge is sent out to all mobile phones in range. The received responses are the SRES and IMSI. These are recorded and used later, along with the COMP 128 algorithm, to derive the shared secret key K that is used.

[174] ETSI (European Telecommunications Standards Institute), http://www.etsi.org/search/frameset/home.htm?CiScope=%2F&CiMaxRecordsPerPage=10&TemplateName=query&CiSort=rank%5Bd%5D&HTMLQueryForm=search.htm&UserRestriction=GSM

[175] Ibid.

[176] Input provided by Rick Fleming, Vice President of Security Operations, Digital Defense Inc.

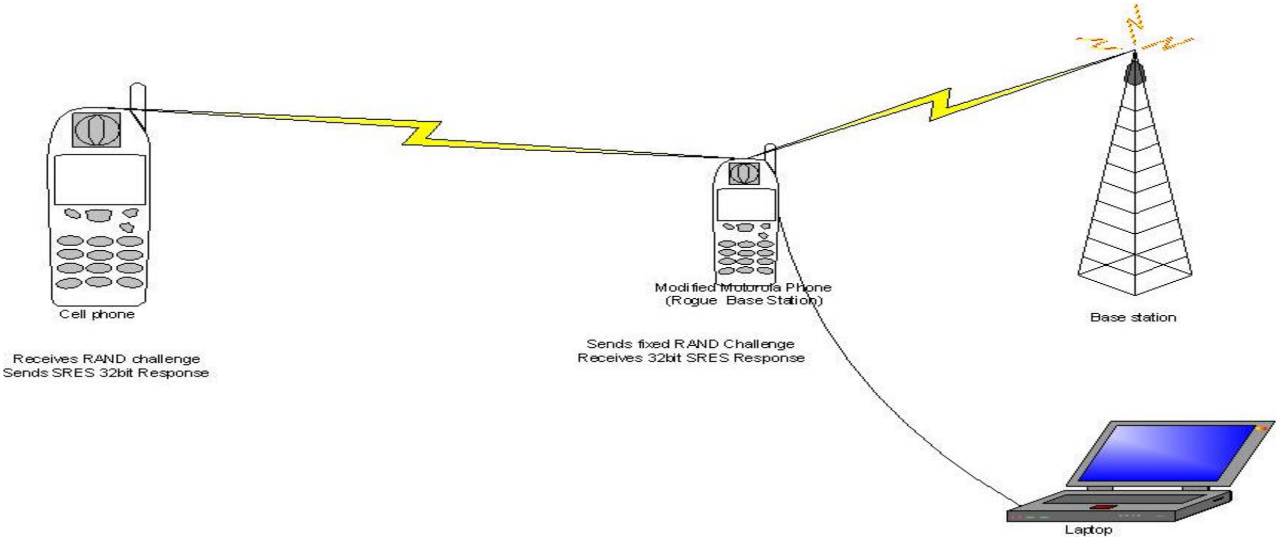
[177] Ibid.

[178] Ibid.

[179] Marc Briceno, GSM Cloning, http://www.isaac.cs.berkley.edu/isaac/gsm-faq.html.

This key is then copied to a smart card and can be used to act as a person or to eavesdrop on a person.

A GSM Hack[180]



The SMS Vulnerability.[181] GSM offers Short Message Services (SMS). SMS is used in GSM systems for many reasons, such as voicemail notification, updating the subscriber's SIM, sending short text messages, and communicating with e-mail gateways. Although these services are convenient, they pose an additional risk to the security of the network. There is freely available software that can spoof SMS messages, send SMS bombs to both handsets and SMS gateways (used to communicate between devices both on and off the network), and corrupt SMS packets that can crash the software on most handsets.

The GPRS Vulnerability. General Packet Radio Service (GPRS) is an IP packet-based service that allows an always-on connection to the Internet. The main problem with this is that it still relies on SMS for Wireless Application Protocol (WAP) push requests. A spoofed (cloned) SMS packet can be sent to the phone requesting a redirected site and fooling users into entering their information into a fake site that they believe is a secure order form. Many GPRS-enabled phones also support Bluetooth, IBM's wireless programming language.[182] Each Bluetooth device has a unique address, allowing users to have some trust in the person at the other end of the transmission. Once this ID is associated with a person, by tracking the unscrambled address sent with each message it is possible to trace individuals and easily log their activities. For Bluetooth devices to communicate, an initialization process uses a PIN for authentication. While some devices will allow you to punch in an ID number, you can also store a PIN in the device's memory or on a hard disk. This is highly problematic if the physical security of the device cannot be guaranteed. Also, most PINs use four digits and half the time they are "0000."

The security of Bluetooth is based on keeping the encryption key a secret shared only between participants in the network. But imagine that you and I are having a conversation using our

[180] Contributed by Rick Fleming, Vice President of Security Operations, Digital Defense Inc.
[181] Ibid.
[182] IBM's wireless programming language.

Bluetooth cell phones. To keep the conversation secure, I use your secret key. Later that day, a friend calls you again and you use your key. Knowing your key, I can use a faked device address, determine the encryption, and listen to your phone conversations. I could also masquerade as you or your friend. Bluetooth authenticates only devices, not users.

WAP Weaknesses. The common flaw in any of these devices, no matter what network, is the Wireless Application Protocol standard, which also includes Wireless Markup Language (WML) and Handheld Device Markup Language (HDML). For the sake of convenience, developers try to require the least amount of keystrokes when entering in credit card number or personal or account information. This means that most of this information is still stored on a server, but the password to access that server is stored in a cookie on the handheld device, requiring only a PIN or sometimes nothing at all to shop online or transfer funds. This means that the actual mechanism used to transport sensitive information end-to-end in these untrusted public cellular networks, is left to Wireless Transport Layer Security (WTLS).[183] Unless 128-bit SSL for mobile commerce or IPSEC for Enterprise access is being used, which most handsets cannot support because they lack processing power and bandwidth, there will be a weak link somewhere in the network that can be exploited. Even then, this only pushes the weakness out to the end devices that are communicating, and it can be easily lost. GSM uses the Wired Application Protocol and also the Wireless Transport Layer Security. This is equal to Secure Socket Layer, but it has weaker encryption algorithms. WTLS is not compatible with SSL, which is the industry standard. Wireless messages travel through a "gateway"[184] that channels them to a wired network for retransmission to their ultimate destination. At the gateway, the WTLS message is converted to SSL. For a few seconds, the message is unencrypted inside the gateway, which in turn makes the communication vulnerable to interception.[185]

Security Solutions for GSM

The inherent problems affecting GSM are not easily corrected. The telephones and PDAs that use GSM technology typically cannot upload protective firmware and software. Users are at the mercy of the telephone developer. Whereas GSM is not vulnerable to war driving like its American counterpart, 802.11, it is suffering from several core vulnerabilities. The 802.11 standard is geared to computers, not handhelds, and thus security can be improved much more drastically for 802.11 than for the GSM protocol. Virtual private networks are the common thread between the two. The establishment of VPNs is commonly referred to as the solution for the existing vulnerabilities of GSM and 802.11. However, when it comes to proper layered security, there are no magic bullets.

To protect information systems that may use any of these technologies, users should deploy virtual private network technology at each and every trusted gateway into their networks and ensure that every user accessing the trusted network uses VPN technology. A virtual private network is essentially a private connection between two machines that sends private data traffic over a shared or public network, the Internet. VPN technology lets an organization securely extend its network services over the Internet to remote users, branch offices, and partner

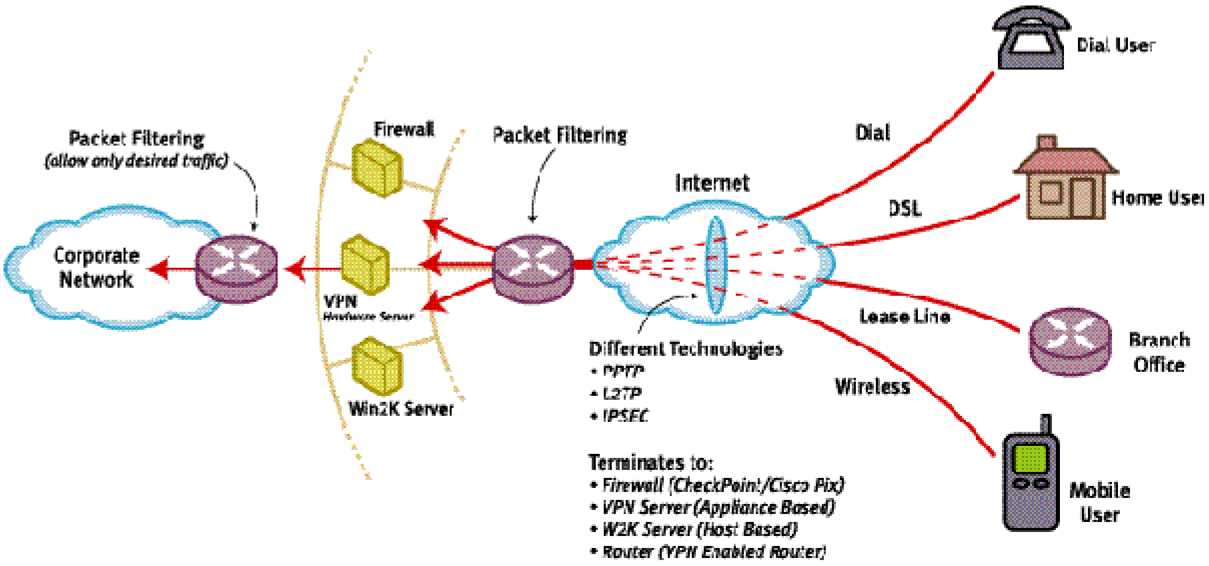
[183] In his paper *Attacks Against the WAP WTLS Protocol,* Saarinen describes in detail a number of security problems with WTLS. Although the WTLS protocol is closely modeled on the well-studied TLS protocol, a number of security problems have been identified with WTLS. These problems include vulnerability to datagram truncation attack, message forgery attack, and a key-search shortcut for some exportable keys.

[184] A gateway is a device that translates the WAP to LAN from wired to fixed-line communication. Hackers have cracked the security for gateways.

[185] Input provided by Dave Thomas, Chief Investigator for the National Infrastructure Protection Center.

companies. In other words, VPNs turn the Internet into a simulated private wide area network (WAN). VPNs allow remote workers to access their companies' servers.

Diagram of Virtual Private Networks



Source: Linda McCarthy, Vice President of Systems Engineering at Recourse Technologies.

To use the Internet as a private WAN, organizations may have to overcome two main hurdles. First, networks often communicate using a variety of protocols; VPNs provide a way to pass non-IP protocols from one network to another. Second, data packets traveling the Internet are transported in clear text. Consequently, anyone who can see Internet traffic can also read the data contained in the packets. This is clearly a problem if banks desire to use the Internet to pass important, confidential business information. VPNs overcome these obstacles by using a strategy called tunneling. Instead of data packets crossing the Internet out in the open, they are first encrypted for security and then encapsulated in an IP package by the VPN and tunneled through the Internet.

Many vendors—Nokia, Cisco, Nortel, Checkpoint, and Microsoft, among others—have viable, secure VPN technologies[186] that can be deployed at multiple locations in a corporate network. While VPNs provide content protection for that information traversing the network, depending on how they are deployed, they may not provide any protection from extraneous users accessing the network itself. In other words, an unauthorized user may not be able to see the content because of the VPN, but that user can still access the network resources and use the bandwidth, causing network congestion and possibly denial of service to authorized users. Access control, authentication, and encryption are vital elements of a secure connection. The Point-to-Point Protocol (PPP) has long been used as the Internet's universal link layer for creating tunnel links

[186] The standards for VPN are currently in revision by the Internet Engineering Task Force to make IP Security more secure, but also to make it compatible with satellite communications.

between devices, but in recent years, the Point-to-Point Tunneling Protocol (PPTP) and Layer 2 Tunneling Protocol (L2TP) have prevailed.[187]

C. A View into the Future: 3G Technology

3G stands for the third generation of wireless communication technology. It refers to pending improvements in wireless data and voice communications through any of a variety of proposed standards. The immediate goal is to raise transmission speeds from 9.5K to 2M bit/sec. In systems and communications security, the goal is not to design a flawless system, but to design a system that can adapt to security enhancements as the need for them is identified. Several of the attacks that were possible on 2G and 2.5G networks have been addressed and eliminated in the 3G environment.

The Strengths of 3G's Security Structure

3G security was based on GSM security, with the following important changes:[188]

- A change was made to defeat the false base station attack. The security mechanisms include a sequence number that ensures that the mobile device can identify the network.
- Key lengths were increased to allow for the possibility of stronger algorithms for encryption and integrity.
- Mechanisms were included to support security within and between networks.
- Security is based within the switch, rather than in the base station as in GSM. Therefore, links are protected between the base station and the switch.
- Integrity mechanisms for the terminal identity have been designed in from the start, whereas they were introduced late into GSM.
- The authentication algorithm has not been defined, but guidance on choice will be given.
- When roaming between networks, such as between a GSM and 3GPP, only the level of protection supported by the smart card will apply. Therefore, a GSM smart card will not be protected against the false base station attack when it is in a 3GPP network.

The 3G system is far more secure than its GSM counterpart. That said, the ingenuity of nefarious individuals should never be underestimated. Certain attacks are theoretically possible on a 3G network. They are described below.

Camping on a False Base Station

This is an attack that requires a modified base station/mobile station (BS/MS) and exploits the weakness that a user can be enticed to camp on a false base station. A false BS/MS can act as a repeater for some time and can relay some requests between the network and the target user but modify or ignore certain service requests or paging messages related to the target user.

The security architecture does not prevent a false BS/MS from relaying messages between the network and the target user, nor does it prevent the false BS/MS from ignoring certain service requests or paging requests. Integrity protection of critical messages may help, though, to prevent

[187] Karen Bannan's article "Safe Passage" in *PC Magazine* August 2001 reviews seven VPN providers for products that would suit a medium-size business with a budget of $10,000 that needed a VPN for its central and branch offices. the article is available at http://www.pcmag.com/print_article/0,3048,a%3D12352,00.asp.

[188] The evaluation of relevant strengths and weaknesses associated with 3G technology was provided by Rick Fleming, Vice President of Security Operations, Digital Defense Inc.

some denial of service attacks, which are induced by modifying certain messages. Again, the denial of service in this case only persists for as long as the attacker is active, unlike the other kinds of attacks, which persist beyond the moment where intervention by the attacker stops. These attacks are comparable to radio jamming, which is very difficult to counteract effectively in any radio system.

Forcing Unencrypted Communications

This attack requires a modified BS/MS. While the target user camps on the false base station, the intruder pages the target user for an incoming call. The user then initiates the call setup procedure, which the intruder allows to occur between the serving network and the target user, modifying the signaling elements so that it appears to the serving network that the target user wants not to enable encryption. After authentication, the intruder cuts the connection with the target user and then uses the connection with the network to make fraudulent calls on the target user's subscription.

Integrity protection of critical signaling messages protects against this attack. More specifically, data authentication and replay inhibition of the connection setup request allows the serving network to verify that the request is legitimate. In addition, periodic integrity-protection messages during a connection help protect against hijacking of unenciphered connections after the initial connection is established. Hijacking the channel between periodic integrity-protection messages is still possible, but this may be of limited use to attackers. In general, connections with ciphering that has been disabled will always be vulnerable to some degree of channel hijacking.

Again it should be pointed out that these attack profiles are theoretical, based on an understanding of how the technology will be deployed. All in all, 3G systems have enhanced and improved security technology in place, but continued vigilance is necessary to maintain their security to set up a mobile-originated call.[189]

[189] Provided by Rick Fleming, Vice President of Security Operations, Digital Defense Inc.

Annex IV. Monetary Transmitters and Internet Service Providers

As a new global money movement industry configuration emerges from the convergence and integration of the telecommunications, computer, and financial services industries, it is essential to identify the integral components of the wholesale and retail payment systems and reassess the regulatory response to this new sector. This annex suggests that by analyzing the form and function of the payment system at large, and the integral role that money transmitters and Internet service providers (ISPs) in particular provide to the industry—whether they mean to or not—ISPs have become critical sectors that can directly affect the safety and the soundness of an institution, as well as the system as a whole, and therefore ISPs must fall under a regulatory net.

This annex delves more deeply into the five basic barriers to creating the new safe and sound environment. These are the lack of common definitions, common regulatory objectives, harmonized regulatory schemes, a shared risk perspective, and common minimum-security prerequisites.

The electronic economy presents a new regulatory paradigm: universal access in a safe and sound environment.

For example, the use of multiple distribution channels to distribute financial services or to make payments has expanded the circle of essential financial service components to include a Web site hosting service, a third party software developer that plans and implements the Web site, application software/service providers, a third party processor to facilitate the movement of information from the Web site to the financial institution's network, a customer service call center, and one or more ISPs, or money transmitters. Increasing the flow of money transmission through these channels means that money is moving at an ever-increasing speed. On the one hand, this increased acceleration enables money to operate more efficiently. On the other hand, it increases risk. In essence, use of these new channels means that the financial services sector now broadcasts; publishes; provides/uses e-mail, Internet services, network services, and entertainment; hosts online forums; and uses bulletin boards. Thus, the very definition of a money transmitter becomes more complex as the classes of nondepository service providers become more varied. It requires a two-part test. One part is the extent to which an institution or the industry relies on that provider to transact and deliver financial services. The other part is the extent to which the provider can affect the payment system.

A. Common Definitions

Toward a Working Definition of a Money Transmitter

Because money transmitters perform a range of services and because until recently they operated in the shadows of the giro payment system, these entities have escaped legislative or regulatory attention. These entities are not well defined, and they are not regulated or supervised in many countries even if they are defined.

Often, money transmitters are referred to as nonbank financial institutions or money services businesses. Numerous definitions exist for this payments system "service" sector. In the context of electronic payment systems, they typically serve as third party ACH providers. Money transmitters may perform a variety of services, including issuance of money orders, wire transfers, currency exchange, check cashing, and check presentment. More recently, they provide electronic check presentment services and point-of-sale money payment order information to the accepting bank. Money transmitters operate outside of the depository institution, but they often are associated in some way with one or more depository institutions in a downstream relationship.

The growth in the range of money transmission or payment services and their potential impact on the larger payment system make it essential to develop a common definition and common regulatory goals and objectives. For purposes of this annex, a money transmitter is defined as any commercial enterprise that engages in the transfer and exchange of monetary instruments and currency.

Money transmitters, however, do not operate alone. Money transmitters require access to the telecommunications system in order to transport the information from point to point. Usually a money transmitter contracts with an ISP to transport the information across network lines.

Defining an ISP

Whether an entity is an ISP can be difficult to determine under existing law. ISPs are not regulated at all in most countries. And countries that have tried to regulate them have seen significant backlash. One recent example is Australia's Broadcasting Services Amendment (Online Services) Act 1999. Often referred to by its critics as the Internet Censorship Act, it has received international attention and is touted as an attempt by one country to impose a censorship regime on the Internet.

The Online Services Act defines an ISP as anyone who provides an Internet carriage service that is used for:

- The carriage of information between two end-users outside the “immediate circle” of the supplier, as defined in the Commonwealth Telecommunications Act of 1997. Where one person uses an Internet carriage service to view the content of a second person (e.g., by visiting a Web site), both of these people would be considered end-users of that carriage service.
- The carriage of information simultaneously to more than one end-user, at least one of whom is outside the immediate circle of the supplier.

Critics decry the act, claiming that it is overly broad. Indeed, a number of entities, including financial services providers, could fall under the act's definition of an ISP. For example, the definition could include all persons or entities operating their own Web servers. Or a business that enters into an agreement with an ISP to provide access to the Internet and that in turn permits other people to access the services provided by that ISP arguably falls under the umbrella of this definition. Companies providing independent contractors with access to the Internet also may fall within this definition. It also would cover specialists that outsource the management of IT services, including Internet access; ancillary institutions that provide Internet access to visitors to the institution; and businesses that allow clients to access their intranets, provided a link exists from the business’s intranet to the Internet and this link is available to clients accessing the intranet.

The act does not provide a meaningful definition for “Internet content host.” It simply states that the term refers to a person who hosts or proposes to host Internet content in Australia. The definition of “Internet content” further fogs the issue. It is defined to mean information available for access using an Internet carriage service. But then it excludes e-mail. Finally, the act does not define which activities constitute “hosting.”

Another confusing example is the European Union directive on electronic commerce (Directive 2000/31/EC of the European Parliament and of the Council of June 8, 2000), which covers certain legal aspects of information society services—in particular electronic commerce—in the internal market. This directive is a legislative standard because it brings ISPs under regulatory authority. But it is an empty-handed gesture because it excludes the ISP from liability except in carefully defined circumstances.

ISPs are not regulated in the United States. In fact, the United States has taken a hands-off approach, letting the industry self-regulate. This is the traditional response in dealing with emerging markets. Historically, the United States does not legislate until blatant and substantial abuses occur, and then typically it enacts only targeted consumer protections.

In section 4, Article 12, the EU directive on electronic commerce proposes that where an intermediary service provider is a mere conduit—an entity that consists only of transmitting across a communications network provided by a recipient of the service, or that provides access to a communications network—that service provider is not liable so long as the provider does not (1) initiate the transmission, (2) select the receiver of the transmission, and (3) select or modify the information contained in the transmission. Neither is the provider liable for the automatic intermediate or temporary storage of the transmission so long as the provider (1) does not modify the information, (2) complies with conditions on access to the information. Moreover, this directive states in Article 14 that hosting providers are not liable if they meet certain conditions in the handling of the information they store.

Unfortunately this directive sets a conflicting and low bar for service providers that are intimately connected to the payments system. Moreover, it evidences a significant lack of understanding about the risk of transacting financial business over the Internet and encourages unsafe and unsound behavior, because in a competitive environment the financial services entity will choose the lowest-cost host.

B. Common Regulatory Objectives

The common regulatory goal is to encourage a safe and sound financial environment by mitigating the risk of payment system failure or institutional failure to the maximum extent possible. Implementing and monitoring electronic security measures is as important to the new financial services configuration as implementing and monitoring financial audit controls was to the old structure. Given this goal, the supporting objectives are to identify, define, and address the core components of the financial services industry and the payment systems from a systems basis and create appropriate regulatory responses to those components.

From a systems perspective, money transmitters and ISPs directly affect the payment system and are integral components of the delivery of financial services in the electronic arena. Because they are essential to the delivery of financial services, their vulnerabilities, if left unchecked, can insert significant risk into the payment system. Therefore, regulators should have adequate assurances that these operations meet common minimum safety and soundness requirements.

The most important objective in a convergent, open, universal access financial services environment is to identify and mitigate operational and systemic IT risks to the fullest extent possible. This dictates that regulators emphasize and work with service providers to identify an institution's operational and systemic IT risks and vulnerabilities. Both regulators and institutions should operate within the following framework:

- Security is a program, not a project.

- Security needs are constantly changing as the system evolves.
- Continual security assessment and risk analysis is critical.
- Operational risk and systemic risk must be eliminated where feasible and mitigated where possible.

Regulators should serve as objective third party monitors and auditors that can assess the adequacy and sufficiency of the institution’s and its service provider's e-risk management approach. Because an entity’s security is only as good as its weakest link, the financial institution must be an active participant in its service provider's e-security program. By the same token, the regulator’s assessment is not complete until it has examined the provider’s risk management approach.

But at present, money transmitters in the United States either are not subject to any regulation or are subject to varying and inconsistent degrees of regulation and oversight. Typically they are undercapitalized, they use little to no risk management analysis, and they operate like mom and pop shops—extremely susceptible to bankruptcy and failure. With the escalation of Internet-related commercial activities, providing countless payment systems conduits has become essential. Thus, money transmitters are escalating the disintermediation of the transitional payments systems.

For example, one e-payment entity provided electronic check presentment services to merchants at the point of sale. The process was similar to a credit card presentment. At the end of the business day the data was batched and routed through the ACH system for presentment and payment. On any given day, several hundred thousand dollars were outstanding. The entity was severely undercapitalized. In fact, the entity was not bonded or insured, and its management had no banking or financial service experience (the president previously owned a construction company). A glitch in the software resulted in a payment backup, which over a 24-hour period caused the entity to go into bankruptcy. The merchants were left with no recourse because the money had been deposited in the entity’s account on presentment and the president had long since disappeared.

As the above example suggests, the primary focus in developing regulatory responses to money transmitters is to enhance measures to deter their use as vehicles for crime. In the example, several investors attempted to finance the venture with proceeds from money-laundering activities. Because money transmitters typically are being used more and more often to launder money, most of the regulatory activity in this area results from anti-money-laundering legislation. Two efforts stand out.

- One is the Uniform Money Services Act, a model act adopted by the National Conference of Commissioners of Uniform State Laws (NCCUSL) in 2000 and known as the Money Transmitters Act.[190] The act requires a money transmitter to obtain a license to operate, sets forth licensing criteria, spells out enforcement and compliance measures, establishes the jurisdiction and scope of the act, and sets forth audit and examination authority. It also contains provisions specifically affecting management; these cover bond, minimum net worth, management experience, and disclosure of prior litigation and criminal prosecution. Only seven states have adopted the act. Most of these state actions became effective in January 2002.
- Before NCCUSL proposed the Money Transmitters Act, the Money Transmitters Regulators Association (MRTA), formed in 1989 as a state regulators organization, had

[190] See www.law.upenn.edu/bll/ulc/moneyserv/UMSA2001Final.htm

created the less comprehensive MRTA Act. It too is a mode l act crafted to deal with the licensing and regulation of money transmitters. As of 2001, regulators in 30 states were members of the MRTA. Only five states had adopted this act by the end of 2001.

C. Harmonized Regulatory Schemes

Until January 2002 money transmitters in the United States were not regulated at the federal level. But with an estimated 200,000 money transmitters operating in the United States, and with mounting evidence that money transmitters are being used to launder money, they have come under increased scrutiny in the last few years. In its 1998/99 Annual Report, the Financial Action Task Force noted a growing trend toward use of nonfinancial professional service providers as conduits for money-laundering and other nefarious activities.

Given that money-laundering and other criminal activity heavily influence this sector, harmonized regulatory schemes that address the criminal risk, as well as the less blatant but equally detrimental operational risk, have become essential. Harmonized regulatory schemes provide a consistent basis from which regulators and law enforcement can use common definitions and common regulatory goals and objectives. In this instance, money transmitters and ISPs that hold themselves out as able to provide services to the financial sector should be regulated and licensed. A license should indicate to the financial services sector that any entity holding it has met the entry barriers. Access to serving the financial services sector should be limited to entities that have the requisite management skills, risk mitigation capabilities, capital, insurance or bonding, and security to mitigate their portion of the risk. Their performance should not increase the risk to the system. And remedies should be provided to those who unknowingly use the services of an entity that does not meet the regulatory requirements.

D. Shared Risk Perspective

Developing appropriate regulatory schemes includes developing an approach to mitigate or manage risk. Here, the concern is that money transmitters and ISPs are not legally liable for the services they provide. Money transmitters and ISPs provide essential services to the new financial service sector. Yet they are not required to post security for their services and they carry no liability. In fact, legislation in some countries holds that ISPs are not liable for transmission failures or losses. Also, because money transmitters and ISPs are not subject to reporting requirements, little information is available on the extent of the vulnerability. But informally the industry acknowledges that these losses occur consistently.

Usually, the money transmitter/ISP venture is structured as a layered relationship. It is built on successive contracts, each containing no liability or very limited liability. The money transmitter provides database software to the end-user. Typically, this software is not warranted or it has limited warranties and the money transmitter carries no liability or limited liability for providing the software or access to the ISP. The ISP typically leases a number of telephone lines or telecommunications resources at a certain rate. The underlying service contract with the telecommunications provider is solely for leased space on the network. The network provider, typically one of the large public switched network providers, provides only a transport mechanism. This arrangement is similar to rights-of-way agreements for utilities or trains, and it is often referred to as “common carriage.” This right-of-way allows one to use the property along the track but does not include access to the track. The ISP contracts with the money transmitter for cost-plus as a transport mechanism only and incurs no liability or limited liability for this service.

The ISP may enter into a service-level agreement with the user (i.e., the money transmitter). Industry standard norms require that the telecommunications system be operational at least 99.5 percent of the time during the service contract. The contract contains a formula for determining an appropriate refund mechanism, dependent on the number of times and amount of time that access falls below the stated service level. The money transmitter in turn provides no liability or limited liability to the user. In addition, the money transmitter provides no additional value in the form of security for its service. The money transmitter simply provides a type of bundled service to the user. In essence, the money transmitter charges a convenience fee. The user simply uses the money transmitter's software to create and store the payment order data, which it then "batches" and sends on an agreed-to periodic basis to a clearinghouse for deposit or credit to the user's account after it has wound its way through the payments system.

Money transmitters and ISPs that hold themselves out as able to provide services to the financial sector should be required by regulation or legislation to provide liability. They should be held to a higher standard of care. Sharing risk is a tried and true model in the financial services arena. Some critics argue that sharing risk would increase the basic service cost, but there is not sufficient information to support this argument. In fact, only when service entities are required to report losses or suspected losses can sufficient information be gathered for either side. It will result in better pricing for bonds and insurance.

The industrial age gave rise to certain concepts by which the telecommunications and financial services industries were regulated. For telecommunications, it was based on public safety, interest, and welfare through the use of universal access and service. And, banking was based on safety and soundness with nondiscriminatory access to credit opportunities

E. Common Minimum Security Prerequisites

At one time, access to the financial system was limited to a chosen few through the use of complex protocols. Today anyone can access the system, using microwave, wireless, satellite, public switched network computer, IP telephony, interactive television, ATM, or the brick and mortar structures. And the financial system is accessible to anyone, anytime, anywhere using cash, debit card, check, credit card, stored value card, credit/debit card, or smart card. Money has become interoperable as telecommunications and computers empower a person to convert money from one currency to another simply by pushing a button. Even the servers of a telecommunications company can be used for the dual purpose of facilitating cellular calls and effecting payments between cell phone subscribers.

Because the financial services sector is the most sophisticated, most dependent, and oldest user of IT, it has been evolving industry security standards for the last three decades. SWIFT is an example of a secure service provider, as is FEDWIRE. Use of these standards needs to be expanded to cover money transmitters in general and financial services ISPs.

Given the new regulatory paradigm, and given the increased outsourcing of operations, the financial regulatory net should be increased to include ISPs and money transmitters as the new core financial service providers. The regulatory framework should be two-pronged. It should involve minimum security standards that the industry itself imposes on its members. And it should involve government-harmonized security guidelines with examination and audit compliance requirements. Requiring entities to report losses or suspected losses to regulators and law enforcement and requiring them to file remediation plans for closing identified vulnerabilities will bolster security. Ensuring that money transmitters and ISPs have adequate security and are shouldering their appropriate share of the risk will enhance the safety and soundness of the new electronic environment.

By setting minimum security standards, enforcing these through appropriate examination and audit controls, grading an institution and its circle of activity's risk management approach and its collective security program, noting whether an institution is reporting losses or suspected losses and filing remediation plans and fixing its identified vulnerabilities, the security of the system as a whole will be enhanced and the benefits will be realized globally. In suggesting the new paradigm, regulators must be poised to define a safe and sound financial environment.