SAND98-0930C

80

SAND--98-0930C
CONF-980433--

SURETY OF THE NATION'S CRITICAL INFRASTRUCTURES: THE CHALLENGE RESTRUCTURING POSES TO THE TELECOMMUNICATIONS SECTOR

Roger Cox, Thomas E. Drennen, Laura Gilliom*, David L. Harris, David M. Kunsman*, Michael J. Skroch
Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-0451

(to whom correspondence should be addressed: lrgilli@sandia.gov, dmkunsm@sandia.gov)*

Abstract

The telecommunications sector plays a pivotal role in the system of increasingly connected and interdependent networks that make up national infrastructure. An assessment of the probable structure and function of the "bit-moving" industry in the twenty-first century must include issues associated with the surety of telecommunications. The term surety, as used here, means confidence in the acceptable behavior of a system in both intended and unintended circumstances. This paper outlines various engineering approaches to surety in systems, generally, and in the telecommunications infrastructure, specifically. It uses the experience and expectations of the telecommunications system of the United States as an example of the global challenges.

The paper examines the principal factors underlying the change to more distributed systems in this sector, assesses surety issues associated with these changes, and suggests several possible strategies for mitigation. It also studies the ramifications of what could happen if this sector became a target for those seeking to compromise a nation's security and economic well being. Experts in this area generally agree that the U. S. telecommunications sector will eventually respond in a way that meets market demands for surety. Questions remain open, however, about confidence in the telecommunications sector and the nation's infrastructure during "unintended" circumstances --- such as those posed by information warfare or by cascading software failures. Resolution of these questions is complicated by the lack of clear accountability of the private and the public sectors for the surety of telecommunications.

Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United Stated Department of Energy under contract DE-AC04-94AL85000.

DISCLAIMER

INTRODUCTION TO THE CONCEPT OF CRITICAL INFRASTRUCTURE

The World Trade Organization recently released a report on the future of the Internet and electronic commerce. It states that there are many technical and legal challenges ahead, which we believe apply to telecommunications in general: "securing adequate telecom infrastructure; providing legal and jurisdictional safety; ensuring security and privacy of information; designing tax regimes; and 'fostering equal opportunity through appropriate policies to promote education and access, particularly in developing countries.'"[Giu] Although not offering solutions, the report identifies two broad considerations that must be addressed in meeting these challenges: governments should not regulate the Internet individually, that any such regulation should be multilateral; and, the issue of equity in electronic commerce among developed and developing countries must be addressed. We assert a third broad consideration: the telecommunication system must be sure. *Surety* of the system encompasses more than the security and privacy issues cited above. System surety is the theme of this paper. Before proceeding to a discussion of issues surrounding telecommunications surety in the new paradigm, we first must present the concept of critical infrastructures and then a general description of surety and the Sandia approach to surety.

The United States relies increasingly upon an infrastructure composed of a complex framework of interdependent networks and systems that provides a continual flow of goods and services essential to the defense and economic security of the United States. Telecommunications is one of these elements; in fact, according to Richard Kuhn of NIST, "The telephone system of the United States is possibly the largest distributed system in existence."[Kuh] Other elements include transportation, electric power, oil and gas distribution, finance and banking, and vital human services. This highly interdependent system of systems is owned and serviced primarily by private industry, but is also an essential element in governmental operations. Our mounting dependence on the infrastructure is, unfortunately, accompanied by increasing vulnerability to failures and disruptions from malicious intrusion, inadvertent error, aging and degradation, natural disasters and, more and more, by the complexity of the system itself as the number of interconnections within and among the infrastructure elements increases. We must address the surety (safety, security, and reliability) of the telecommunications infrastructure and those elements of the global infrastructure that affect national security.

Many elements of critical infrastructures are vulnerable to both physical and cyber threats. Some cyber threats can be mounted at very little cost to an adversary and are hard to detect and trace. In the banking and finance community, for example, consequences of a successful cyber attack could be staggering. Trillions of dollars in funds and securities are transferred daily by electronic communications mechanisms. The high value and volume of such transactions within an open environment exposes the business community and its customers to severe potential consequences from accidental or deliberate alteration, substitution, or destruction of data. This potential is compounded by interconnected networks and by the increased number and sophistication of malicious adversaries (see, for example, [Ric] and [Fis]).

The interconnection between elements of the infrastructure increases its vulnerability on a national level. If designed and implemented properly, the interconnection could reduce the number and severity of failures; however, without adequate forethought, the coupling may cascade failures within and across diverse elements of the infrastructure. As the infrastructure systems of telecommunications, electric transmission, and gas transmission change topology and technology in the face of deregulation, the system could evolve so that it is more vulnerable to intentional or unintentional disruption.

Even though terrorist attacks on pieces of the critical infrastructure are not as common in the U.S. as they are in Europe and Latin America [OTA], recent events within the US such as the Oklahoma City and New York Trade Center bombings illustrate that these statistics may change. Publicly available documentation contains sufficient information to identify potential power grid targets that could result in major, long-term blackouts.

Historically, private industry has not addressed certain low probability of occurence but nevertheless high-consequence events because of the high potential cost of protective measures and the lack of a corresponding clearly defined benefit from those measures. However, as interdependency between infrastructure elements continues to build, particularly between the banking/finance, telecommunications and electric power industries, some additional steps may have to be taken to understand and predict how national security and the economy might be affected by potential cascading impacts.

Introduction to the Concept of Surety

The term "surety" may be unfamiliar to the reader. "Surety" was first used in a military context and included readiness, reliability, and other similar issues. The term was adopted by the United States Atomic Energy Commission (now the Department of Energy) and then its laboratories, including Sandia National Laboratories. Surety is an attribute of a system, one of the most important attributes. We define surety as the confidence that a system will perform in acceptable ways in both intended and unintended circumstances. The latter include abnormal (accident) and malevolent (intelligent attack) situations. Elements of surety include reliable performance, safety, and security. Within the security element are three sub-elements: protection against unauthorized access, protection against unauthorized use, and permission for authorized access and use. The three high-level elements of reliable performance, safety, and security are not orthogonal. In fact, they are closely associated; unreliability can make a system less safe and too many safety and/or security features can make a system less reliable or less safe.

A sure system, then, is a system in which, and through which, significant events occur if, and only if, the designer and authorized operator intend them to occur.

Approaches to System Surety

Two primary theories have been proposed for achieving surety in high-consequence systems: a theory based on high-reliability and one based on normal accidents. For example, John von Neumann asked whether it is possible to build reliable systems from unreliable parts. For high-reliability theorists, the answer is "yes," and this is achieved in part through appropriate system design and operations. Some high-reliability theorists advocate anticipation and resilience-- the former is to design a system that prevents undesired events from occurring and the latter is to design a system that mitigates the effects of the undesired event once it has occurred. Scott Sagan, a Stanford University political scientist with expertise in organization theory, writes that high reliability theorists contend prevention and mitigation are enhanced by the application of redundancy, and that in addition, they "...believe that hazardous technologies can be safely controlled by complex organizations if wise design and management techniques are followed." [Sag] On the other hand, Diane Vaughan's recent study of the Challenger accident found that following established procedures in a system with many engineered redundancies is insufficient to assure surety if organizational structure and processes are flawed.[Vau] Management structures and processes can have as many complexities and couplings as any hardware or software system.

In contrasting the two theories Sagan writes, "[Normal accident] theory predicts that serious accidents are inevitable if the organizations that control hazardous technologies display both high interactive complexity (which produces bizarre and unanticipated failures) and tight coupling (which causes the failures to escalate rapidly out of control)."[Sag]

Charles Perrow coined the term "normal accident." He writes that "Most high-risk systems have some special characteristics, beyond their toxic or explosive or genetic dangers, that make accidents in them inevitable, even 'normal.' This has to do with the way failures can interact and the way the system is tied together....The odd term *normal accident* is meant to signal that, given the system characteristics, multiple and unexpected interactions of failures are inevitable. This is an expression of an integral characteristic of the system, not a statement of frequency."[Per, emphasis in the original]

Several systems-engineering approaches exist to achieve a sure system. They are not mutually exclusive, although proponents of the two prevalent academic theories would likely disagree. In the design and operation phase, the approaches can complement each other. The approaches involve risk management, formal management of large, integrated, complex systems, and reliance on first principles to achieve a sure system. Another and more usual approach is a hybrid of the other approaches, taking certain elements of each and combining them. In any design for surety, organizational factors must be evaluated in terms of how they affect surety of the system. We discuss the merits of these several approaches below.

Risk management for high-consequence technologies may well have been born with the seminal paper written in 1969 by Chauncey Starr, now president-emeritus of the Electric Power Research Institute [Sta, 1969]. He posited then that the performance criteria for a mature technological system is a balanced trade-off of societal benefits and societal costs, indirect as well as direct. He more recently returned to these issues: "In a world in which zero risk is impossible, and safety has no absolute measure, the question then arises as to how to determine 'how safe is safe enough."[Sta, 1991] (In keeping with the three elements of system surety, we could also ask, "How reliable is reliable enough?" and "How secure is secure enough?")

Risk management explicitly recognizes that off-normal situations will arise and seeks to manage the risks associated with those situations. It is risk **management**, not risk **elimination**.

Risk management starts with an objective assessment of the risks and benefits of technologies and their applications. Examples of risk assessments abound; for example, see [USN]. The aim of risk management is to provide a rational basis for decision making, and such analyses should include uncertainties. Risk management techniques have also been applied to the analysis of the regulation of risk [Har].

The use of formal methods in system design. The literature on systems engineering is almost unanimous in recognizing the existence of major difficulties in the specification and design of large, integrated, complex systems. Always a problem, this is becoming increasingly acute with the use of microprocessors that are evolving to ever-smaller sizes with ever more possible components and interconnections. Simply developing and maintaining an accurate characterization of a large carrier network is a major challenge. Verifying and validating the integrity of the states that can occur for all network elements and services across carriers would be a daunting task, as it also is to verify and validate the integrity of the states that occur in the evolving telecommunications network. After all, as mentioned above, the telephone system of the U.S. is huge and is becoming both larger and more complex on almost a daily basis.

A recent paper concerning formal methods notes, "A [specific] problem encountered when designing high-consequence, high-assurance systems involves using unclear, informal, and non-rigorous techniques to design and analyze the system. This problem manifests itself in the systems-level design when describing reactive behaviors." [Dav] (Reactive systems continuously react to both external and internal stimuli.)

More formal methods do exist. "Robust product design" is one that uses quality approaches to optimize design performance [Fow]. Another, championed at Vanderbilt University, involves paradigm-concept development. It borrows aspects from analytical and model generation concepts developed in computer science, in particular object-oriented analysis and finite-state machines. Risk assessment analysis techniques such as fault tree analysis or failure modes and effects analysis have proven problemmatic for examining systems with complex timing dependencies. The newer computer science-based approaches should better address this limitation.

First Principle Approach. Instead of attempting to assess and manage risks or create intricate state machines, the system designer can rely on first principles to achieve a sure system. For example, newer nuclear reactor designs rely on physics to provide inherent safety; in some such designs, water cools the reactor core in an emergency by being fed by gravity from a tank through a pipe with a fusible drop-out plate in it which melts at accident temperatures.

We at Sandia use first principles in nuclear weapon safety so that we do not need to analyze every possible perturbation of accident environment using risk assessment - or other techniques. Instead a safety theme for a weapon is developed, generally using the "three *I*'s." [D'An, Plu] *Isolation* of critical components from their surroundings; assuring that a firing signal is *incompatible* with any other that could be generated in either normal or abnormal environments; and *inoperability* of the system if certain abnormal circumstances occur. It can be shown mathematically that, in theory, this first principle design approach is perfect [Kun]; engineering implementation, however, may be less so. Proper design, manufacture, and assembly can create a predictably safe high consequence system.

The design theme of the three *I*'s was specifically created for the safety element of a system. *Isolation*, *Inoperability*, and *Incompatibility* can also be extended into the security element, but security themes are less well articulated. Recently, some at Sandia have labeled this the three *D*'s for security: *delay*, *disable*, and *discriminate* [Tal]. Others keep the three *I*'s but define them somewhat differently for security than for safety [D'An]. The concept of *delay* for an intelligent attack against a system is to make penetrating the system difficult, if not impossible. Included is the concept of detection, that the authorized operators of the system know they are being attacked. In addition, one time boundary of delay is eternity, that is, a limiting situation is defeat of the attack and the system is not penetrated.

Even if the designers and authorized operators believe that the system will defeat all attackers, if the failure of the system would cause high consequences, it would still be prudent to design and operate the system with the second *D* in mind. If the system is penetrated, the attacker becomes the unauthorized user who will disrupt the normal operation of the system by either ceasing its normal operation or causing the

system to perform operations that are not authorized. Designing elements of the system to *disable* itself if operator authentication is lacking is the second *D*.

For a telecommunications network, temporary, localized disablement is the preferred feature rather than a more severe disruption of the system. Care must be made in designing the reversibility, so the attacker does not automatically gain access to it when he penetrates the system.

Lastly, the system must be able to *discriminate* between an authentic message and a non-authentic one. Cryptographic methods can be used. Also included in this third *D* is the concept of thwarting an attack whose purpose is to deny authorized use.

Hybrid approaches are more common in system design. Rarely does the system designer or analyst use just one approach to achieve or assess system surety. It is certainly prudent to use nature to achieve surety as much as possible, but engineering designs are not theories, and implementation of a design occurs in an imperfect human world. Hence, hybrid approaches are generally used. It is also prudent to simplify and decouple the system as much as possible so that fault propagation is limited and mitigation strategies can be used. More complex is not necessarily better, nor is more tightly coupled necessarily more effective when all aspects of surety are considered.

How organizational factors affect surety must also be considered. [PatC, Kai, Sag, Vau, Zeb1989, Zeb1991] Not just hardware and software are important, but also the structure of and communication within the organization designing and operating the system. "Accidental experiments"[1] of the last two decades include Three Mile Island, Bhopal, Challenger, Chernobyl, Piper Alpha, and airline crashes such as that of American Airlines at Cali, Columbia. Some findings to date, with 20-20 hindsight, are that all involved human error, questionable operations, design flaws, and most involved problems at interfaces. But, to a large extent, these do not seem to be the root causes. Rather the root causes were the systems themselves, which were complex and tightly coupled, often with unforeseen interactions, in spite of much testing and analysis. The systems were unforgiving, and, in general, the organizations did not approach them with a broad examination of all possible system states and often kept responsibilities compartmentalized at too high a level. In addition, in some cases, the accident that happened could not have happened according to the understanding of the designers and the predictions of their models. In a majority of cases, the accidents against which the design was analyzed were not among the type of accidents that were actually experienced.

The lessons are: (1) to think broadly for the definition of the system to include operations, normal and off-normal, and operators, humans and procedures, as well as hardware and software; (2) to avoid compartmentalizing aspects of the system at too high a level; (3) to have someone assessing, continually, the entire system; and (4) to make the system fault-tolerant to the real threats it may encounter. In a highly competitive marketplace with a large number of new entrants, we would expect an intensification of compartmentalization to occur at the level of individual carrier networks. The informal cooperation that may have developed among a few large carriers to manage widespread failure events will be challenged in the future. Cost pressures and the lack of trained staff may make the dedication of resources to simply monitor the network too expensive for many new entrants, and perhaps even for more established providers.

TELECOMMUNICATIONS

Description

We include in the definition of telecommunications not only telephony but also a wide variety of communications media and devices. For the U. S., the Communications Act of 1934, as amended by Congress in 1996 provides the following definition:

> The term "telecommunications" means the transmission, between or among points specified by the user, of information of the user's choosing, without change in the form or content of the information as sent and received. [Act]

[1] This phrase was coined by Henry Petroski of Duke University.[Wal]

A key phrase in this definition is "information of the user's choosing". This encompasses the wide spectrum of information types and media types that current technology allows and which technology is envisioned to accommodate. Telephones are now used for both voice and data communications, data that may be facsimile information, digital data that uses a telephone connected to a computer or computer network, and others. Cellular, satellite, and other wireless telephone systems are now being widely used as well and allow communications to take place from varying locations or from remote areas. This definition does not restrict telecommunication to current technology. It is technology independent allowing for newer technologies as they are discovered and developed. The impacts of the fact that technology has advanced to provide new operational concepts and new ways to deliver information are only now beginning to be realized. For example, the advent of on-demand-information content opens up new service possibilities.

With the advent of the "digital age", information to be transmitted may be of a variety of types: audio, video, facsimile, textual, and the many other forms in which information can be communicated. Each of these media types use the same underlying telecommunications infrastructure with each type being converted into digital bits, transmitted among points regardless of their content, and when received are presented in the intended form. The intended form is no longer limited to the form in which the information was created. Content can be presented in different ways determined by either the sender or the receiver of the information.

The telecommunication infrastructure necessary to support the transmission of this wide variety of information consists of a vast network of heterogeneous equipment and their integrated software systems with each element of this infrastructure contributing to the provision of the collection of telecommunication services. The complexity of this system defies its comprehension by any individual. One author states that "[a] highly intercommunicating decentralized structure shows far more resilience and likelihood of survival. It is certainly more sustainable and likely to evolve over time."[Neg] Information within this system is transmitted over millions of miles of copper, coaxial and fiber cables, and includes satellite and ground based wireless facilities as well. The switching of information among points specified by the user is a key element of the infrastructure and is typically accomplished with automated, computer-controlled switching equipment. The systems within the telecommunication infrastructure provide for management of the infrastructure as well as its administration and operation. In essence, the telecommunications infrastructure moves bits around the world for the benefit of and at the request of the users.

The Act also states that the Federal Communications Commission shall be constituted for the purpose of regulating communication commerce to provide:

> a rapid, efficient, Nation-wide, and world-wide wire and radio communication service with adequate facilities at reasonable charges, for the purpose of the national defense, for the purpose of promoting safety of life and property through the use of wire and radio communication [Act]

Policy and Business Drivers to a More Distributed Future

Changes in technology and the implied economic benefits of a more competitive industry have been the fundamental drivers behind deregulation of the telecommunication industry. From a historical perspective, the telecommunications industry, initially the telegraph and later the telephone, were regulated because it fit the classic definition of a natural monopoly. Allowing competing services would have led to duplication of infrastructure in the form of duplicated telephone lines and switching equipment.

Over time, technological change and the potential for large consumer benefits eroded the logic behind maintaining natural monopolies and led to changes in the existing regulations. For example, in 1963, Microwave Communications Inc. (MCI) filed applications with the FCC to construct and operate microwave service from St. Louis to Chicago, thereby bypassing the existing common carrier. These carriers opposed the proposal on the grounds that it would result in duplicative, and hence inefficient, services and that MCI was financially and technically unqualified to operate such a system. [Phi, p. 766] The FCC voted to grant the permit, ultimately noting that "...it would be inconsistent with the public interest to deny MCI's applications." [Phi, p. 766] Within 12 months of this ruling, 37 additional applications were filed, proposing the construction of some 1713 microwave stations. By 1971, the FCC had officially announced a policy of free access by new specialized carriers that could meet financial and technical standards.[Phi, p. 766]

As technology continued evolving, additional challenges to the traditional role of the common carriers arose. And over time, the common carriers had greater and greater difficulty convincing the FCC that justification remained for protecting them from competition. Kestenbaum concludes that by 1971 the FCC had adopted a legal-economic judgment that "...reliance on regulation [should be] a last resort, justified only when competition is not feasible or practical."[Phi, p. 772]

While the changing technology made it possible for new companies to challenge the common carriers, the FCC allowed increased competition because of the potential for large economic benefits to consumers. A study by the Crandall and Ellig [Cra] of the economic benefits of deregulation in several industries supports the view that deregulation benefits the consumer in terms of lower relative prices. Crandall and Ellig estimated that deregulation led to real price reductions after ten years of 27-57% in the gas industry, 40-47% in the long distance telecommunication industry, 29% in the airline industry, 28-56% in the trucking industry, and 44% in the railroad industry.[Cra, p. 2]

By the time that the Telecommunications Act of 1996 became law, technology had further blurred the artificial distinctions between various telecommunication providers. Specifically, as the use of digital technologies replaced analog technologies, the distinction between cable, computer, utility, and telephone companies as potential telecommunication providers became less clear. Telephone companies could deliver movies, radio broadcasts, or newspapers on demand via its lines directly to the user's TV or computer when they want them, rather than at preset times. Utilities could electronically read meters and control the flow of electricity to houses based on time-of-day pricing.

These distinctions between telecommunication providers will continue to evolve in the future. Traditionally, one of the FCC's roles has been to allocate the spectrum, so that radio, wireless telephone, and television signals are broadcast in such a way as to not interfere or overlap with each other. However, as digital technologies have evolved, the distinction has blurred. Consumers can benefit in choosing how to use these digital flows. Negroponte [Neg] uses the example of forecasting weather. Under the old formula, a television station might use its share of the spectrum to broadcast a weather report to your television. Technology now allows the broadcaster to transmit the weather in bit form to your home personal computer/TV that then, based on your preferences, displays the weather in traditional form, or as a printed map, or as an individual newspaper update, or as a voice report. "The broadcaster does not even know what the bits will turn into: video, audio, or print. You decide that."[Neg, p. 55] This example illustrates that the technology will continue pushing the regulation over time; the current regulatory system, even after the Telecommunications Act of 1996, is still not able to deal with the digital future. As technology and applications advance, the regulatory structure will have to continually change to catch up; but, however it changes, the structure must respond quickly. A recent editorial in *The Economist* stated "There are clear lessons from [British Telecom's recent] experience for both regulators and other embattled incumbents. For regulators, it is that shock treatment is best. Get out as soon as telecoms can be treated like a normal competitive market—the technology will make it happen sooner than you dared hope."[Eco]

Technology and economics are the ultimate drivers behind telecommunication restructuring. Restructuring has led to a more decentralized, distributed, and less well-defined telecommunications industry. The FCC's role and responsibilities will continue to change as the definition of the "public interest" changes. Future roles for the FCC include enabling competition, maintaining overall system integrity and reliability, and settling disputes among various competitors. Its job will be less and less to determine how the broadcast spectrum is used (since technology advances should continue to relieve bandwidth limitations) and more on how to establish standards for competition. The future, though, will probably not be totally deregulated; regulation to some degree will still be necessary. As Cairncross has observed, "This business rewards size. The more people and businesses a network connects, the greater the value of being plugged into it. Without regulation to sustain competition, the telephone network might naturally revert to a single giant. Even if services remained competitive, the network probably would not...So regulation is probably inevitable to ensure competition."[Cai]

Other Public Impacts of a More Distributed Future

The external public impacts are many, and most concern cultural changes, which we are not qualified to address. We do not know the unintended cultural consequences of the telecommunications revolution. Arthur Schlesinger, writing in *Foreign Affairs*, considered the effect of telecommunications interconnectivity and rapidity on the body politic. "...[T]he interactivity introduced by the Computer

Revolution makes "pure democracy" technically feasible on a national scale....The rise of public opinion polls, focus groups and referendums suggests popular demand for a finished democracy. With a nation of computers plugged into information and communication networks, "full democracy" is just around the corner....[I]s this a desirable prospect?" [Schl] He argues that it is not, that "interactivity encourages instant responses, discourages second thoughts, and offers outlets for demagoguery, egomania, insult, and hate."

Our intent is not to debate the political and cultural ramifications of telecommunicative citizenry except from one specific viewpoint. If such voting occurred, would the integrity of the result be unquestioned? How would the sanctity of an individual's ballot be assured? If telecommunicative citizenry happens, loss of confidence in the process might have a devastating effect on the body politic with even a greater decline in the public's respect for institutions. Sure telecommunications for this application are a must.

Faster telecommunications and larger, more accessible databases can also challenge the privacy of individuals. This has always been a concern whenever databases have been kept, but the future of telecommunications increases the concern by a significant degree. Already today we see information "brokers" on the Worldwide Web offering to access salary information on an individual for $75, telephone records for $80-200, and ten year medical histories for $400 per person [Ber]. Either the public will come to accept this loss of privacy or it will rebel against such an invasion. Privacy standards and regulation, whether by the business community or governments, may be preferable.

THREATS TO THE SURETY OF TELECOMMUNICATIONS

One way to look at threats for telecommunications in the future is to separate internal from external threats: internal threats are based on the technologies used and the organizational structures using them, and external threats are from outside individuals or groups intent on harm. We briefly explore them in this section with the emphasis on those threats external to the system. Threats due to external but natural phenomena (e.g., weather) should be noted but are not further discussed. Threats posed by malevolent insiders are also discussed below.

Internal System Threats

Without proper attention, surety may become more and more difficult to achieve, as the pace of innovation in telecommunication increases from its own recent, fast acceleration. This is the internal technological threat. An article in the *London Times* last autumn discussed this issue:

> Unfortunately technological churning also prevents improvements in reliability and quality. Would you feel safe in an aircraft if you knew that Rolls-Royce and Pratt and Whitney were reinventing the jet engine every six months?
>
> Technological churning also prevents the establishment of common standards that would make equipment compatible. [Kal]

In the old telecommunications world, the public took reliability, quality, and compatibility for granted. The challenge for the future is to assure continuity in surety, so that the changes wrought by technology are as seamless as possible. For example, the numbers needed to ring, dial, or punch a particular person's telephone may change in their digits, but conversation will still occur reliably and with high quality, whether analog or digital, and any other bits transferred through the connection will not be routed wrongly nor not at all. In addition, we do not want any fault, when one occurs, to propagate throughout the system, bringing it all, or a substantial part of it, to a halt. As Birman writes, "Why do distributed systems crash? If we exclude systems that fail because they were mismanaged or poorly designed, the most common scenario involves an isolated problem at one site that triggers a chain of events in which program after program throughout the network eventually shuts down."[Bir] Some believe that the 1990 nation-wide AT&T failure was due to the complexity of the network and its potential for cascading failures.[Kuh] Will the new deregulated telecommunications system be more or less susceptible to such failures? In a later section, we present some ideas on how to approach designing a system so that it will be less susceptible.

External System Threats

A second class of challenges to the future telecommunications world is posed externally, by persons intent on harm, and arises because of changes to society at large. We introduce the changing external threat picture in this section and, further below, discuss what that threat can do to the telecommunication infrastructure in a section entitled "Information Warfare." The external threat can be posed by any number of people-- hacker, terrorist, rogue business, rogue country, or organized crime. The intentions of all, though, are surreptitious information gathering, data manipulation, or system malfunction. Walter Laquer, of the Center for Strategic and International Studies, has written

> Society has also become vulnerable to a new kind of terrorism in which the destructive power of both the individual terrorist and terrorism as a tactic are infinitely greater. Earlier terrorists could kill kings or high officials, but others only too eager to inherit their mantle quickly stepped in. The advanced societies of today are more dependent every day on the electronic storage, retrieval, analysis, and transmission of information. Defense, the police, banking, trade, transportation, scientific work, and a large percentage of the government's and the private sector's transactions are on-line. That exposes enormous vital areas of national life to mischief or sabotage by any computer hacker, and concerted sabotage could render a country unable to function.
>
> An unnamed U.S. intelligence official has boasted that with $1 billion and 20 capable hackers, he could shut down America. What he could achieve, a terrorist could too. There is little secrecy in the wired society, and protective measures have proved of limited value: teenage hackers have penetrated highly secret systems in every field. The possibilities for creating chaos are almost unlimited even now, and vulnerability will almost certainly increase. Terrorists' targets will change: Why assassinate a politician or indiscriminately kill people when an attack on electronic switching will produce far more dramatic and lasting results?...If the new terrorism directs its energies toward information warfare, its destructive power will be exponentially greater than it wielded in the past—greater even than it would be with biological and chemical weapons.[Laq]

Although some of the statements in this example are not strictly accurate, the general notions underlying the conclusions are worth considering -- New telecommunications technology makes it easier for the terrorist to communicate demands to the public since they cannot be as easily isolated from the media. But ironically, the same technology enables the terrorist to have relatively secure communications with their fellows and to perform telecommunications attack from afar.

As another recent article noted, "The future economic terrorist could attack U.S. corporations through their dependence on information. Their use of Trojan horses, viruses, and spoofing could render information systems inoperative. Proprietary concepts and designs could be stolen, test results altered, research and development derailed, damaging data exposed, or personnel records modified."[Med] There is also no editor on the Internet. An attack could simply be the spread of malicious gossip about an individual.

For all the fears of a terrorist assault on telecommunications, however, some fear assaults by others more. Laquer writes, "...the vulnerability of states and societies will be of less interest to terrorists than to ordinary criminals and organized crime, disgruntled employees of big corporations, and of course, spies and hostile governments." [Laq] Furthermore, Medd and Goldstein note that the states that the U.S. labeled as sponsors of terrorism in 1994 (Iran, Libya, Cuba, Syria, North Korea, Iraq, and Sudan) had a combined GNP less than one-fourth that of organized crime! [Med]

How likely are external threats? They are occurring now. Among other examples, the Cali drug cartel has engaged in electronic terrorism against the Colombian government [Med] and drug gangs in Amsterdam have as well against the Dutch police [Rath]. In more direct extortion, since 1993, City of London and New York financial institutions have been attacked forty times by cyber criminals who were able to extort $600 million [Rat a]. The accuracy of these reports is perhaps not as important an issue as the credibility of the threat.

As to terrorism, its frequency fluctuates with the regional and global political climate. The annual number of terrorist incidents against the U.S. has generally decreased over the past three decades (since the

State Department and F.B.I. began keeping such statistics) but the severity of the average incident per year (measured by average number of casualties per incident) has increased. Since the early 1980s, businesses have been the predominant target for terrorism when U.S. public or private assets are the targets, and this trend is growing. Medd and Goldstein noted that by 1994, "...business targets dominate the statistics by nearly a factor of two over the next closest U.S. target type..."[Med]. Furthermore, they argue that organized crime connections with terrorists will increase and that economic motives will continue to be most important. In the 1960s and 1970s, the motive for terrorism was mainly political; in the 1980s, it was religious. Now and in the future, the motive is often, and will likely often be, economic.

A political solution may become much more difficult. Deterrence against, and boycotts of, countries that sponsor terrorism may become less effective than now because that sponsorship is also changing. The financial backing for many terrorists comes not from a state but from wealthy individuals. [Ger, Mac]

ISSUES OF A MORE DISTRIBUTED FUTURE

This section briefly reviews the pre-competitive telecommunications infrastructure. The next subsection then discusses the changes brought on by competition. In general, today's infrastructure is more robust, and service offerings are more diverse. Competition does present new challenges to surety. Surety may be critical in the future in an environment of more distributed technology and more sophisticated threats.

The Pre-Competitive Infrastructure

The pre-1984 telecommunications infrastructure consisted primarily of large regulated monopolies, such as AT&T and the Bell System in the United States and governmental entities (so-called PTTs) in Europe and Japan. Although those networks were quite large, they had fairly simple service requirements by today's standards. First, they were optimized for one service – 4 KHz voice. Other services, such as television, data and fax, either used separate networks, such as cable television, or operated within that 4 KHz constraint, like the FAX. Second, cellular telephones were not yet widely used. Thus, customers had fixed locations that used well-characterized wired media. That basic service mix greatly simplified traffic engineering.

As noted above, the number of service providers, the common carriers, was limited, further simplifying the system. In addition, those providers usually bought proprietary hardware and software systems from a limited, and often vertically integrated, vendor base. In the United States, an example was Western Electric, the captive supplier of the Bell System.. So, although the equipment was nominally standards-based, those standards were often just pro forma publication of the existing, well-tested interfaces of the captive-suppliers. Although that limited supplier base did simplify problems of interoperability, in retrospect, it also led to increased cost and a slow pace of innovation. Finally, the limited number of service providers also allowed a closed signaling network, which increased surety, once out-of-band signaling was introduced.

The regulation of the business placed the complexity, and hence cost, into the switching and transmission elements that led to relatively low-cost and reliable equipment for the customers. This business approach also severely limited the types and brands of equipment that could use the network. So, while the pre-competitive infrastructure did provide reliable, universal service, it also stifled service innovation. At the same time, the regulated monopoly provided, in essence, lifetime employment for the maintenance and operations staff. While this produced a highly-skilled and loyal workforce, it also helped increase service costs.

The twin deficiencies of decreased innovation and increased cost forced the gradual introduction of competitive elements into the telecommunications infrastructure. That process began in the U.S. during the 1970s. In other countries, telecommunications remains a regulated monopoly. The next subsection details some of the infrastructure changes and their effect on overall network surety.

The Competitive Telecommunications Infrastructure And Its Implications for Surety

Multiple Unregulated Service Providers. The present competitive infrastructure differs from the older telecommunications infrastructure in several important ways. The obvious change is the move to multiple,

unregulated service providers. The competition among them has reduced telecommunications costs. It has also presented the potential for increased reliability in that a cautious business can split its traffic among different providers. Other factors may tend to offset the potential for increased reliability, however. For example, calls now are commonly routed through more than one service-provider network, which can increase the difficulty of determining the exact nature of a problem should one occur.

Standards-Based Networking. The new competitive service providers usually do not have captive suppliers. (Even AT&T finally divested itself of its manufacturing arm, now Lucent Technologies, in 1996.) As competitive entities, those new service providers must have competitive equipment prices from their vendors. This often drives them to a hardware and software mix from multiple vendors which means that interoperability is now driven by vendor compliance with industry and international standards. Time-to-market concerns with this quickly changing technology usually preclude producing the highest quality standards. (Version 1.0 of any standard is often filled with bugs.) Another unintended surety-related result of standards-based networking is the loss of security via obscurity. Industry-standard software platforms, such as UNIX and Windows NT, have well-known vulnerabilities. In addition, smaller service providers often use freeware, such as Linux, for which the source code is freely available. In that case, new vulnerabilities are easier to find since adversaries have access to the underlying source code as well. To offset this potential disadvantage, the source code is also available to more defenders capable of working as a distributed community to improve its security. For example, when the SYN floods appeared in recent years, Linux patches were available within a matter of hours, while many vendors didn't release patches for days or even weeks.

Multimedia Traffic Mix. Modern networks must carry multimedia traffic (voice, video, real-time data and non-real-time data) in one integrated network fabric. This complicates traffic engineering. Overloads currently cause about half of the outage-minutes in the U.S. telephone system [Kuh]. The overloads occur when the actual service demand exceeds the capacity of the network. For the older infrastructure in such situations, the system was designed to block a specified small fraction of the call-origination requests. Erlang first modeled that blocking probability in the early 1900s. (Indeed, the Erlang-B and Erlang-C models [Kle] are still useful for rough-cut voice-traffic planning.) Subsequent decades of research have refined those models. From a business perspective, well-proven, single-service models allow the monopoly executives to justify their rate requests to their regulators. From a surety perspective, good models help guarantee network service levels.

The case of multimedia is considerably more complex [Wir] for at least three reasons. First, there are new traffic types in addition to voice. Some multimedia types are minor extensions to the 4KHz voice transmission. For example, constant bit-rate (CBR) compressed video, such as MPEG-1 and MPEG-2, can still use the classical traffic-engineering techniques. Most new services are "bursty," however, because their peak and average bit-rates differ. One example is web browsing. Users may spend several minutes reading each web-page but then want the next web-page to arrive in seconds. Traditional models assumed that this bursty behavior disappeared if call statistics were aggregated over a large enough user population. Recent research indicates that assumption is false [Lel], and new models need to be developed. The second reason is the traffic mix within each multimedia call. The call origination times and call-holding times for both residential and business voice-calls are well characterized. Multimedia calls are more complex. For example, a business call may start as voice and then become a video-conference with occasional file-transfers. Hence, even if each multimedia application has well-known traffic characteristics, the overall traffic characteristics of the call may still be difficult to predict. Furthermore, the increased rate of service innovation makes modeling even more difficult because the traffic mix may change rapidly. The third reason is the growth of wireless and mobile networks. The next subsection discusses that topic further.

Wireless and Mobile Networking. Wireless connectivity and user mobility are undergoing explosive growth in the 1990's. Indeed, in some developing countries, the cellular network is being built before the wired infrastructure. In addition, the cellular service providers are migrating from first generation analog cellular systems, such as AMPS [Rap], to second generation digital systems such as GSM [Meh]. Furthermore, several Low Earth Orbit (LEO) satellite systems, such as Teledisc and Iridium, should become operational in the next few years. Finally, wireless LANs are appearing in factories, warehouses and offices. So, for a

price, users will eventually have seamless, worldwide access to their preferred communications and computing environments.

These factors have a mixed effect on network surety. Mobility certainly complicates traffic engineering. It also increases the signaling load on the network. In older networks, users usually signal only during call origination. In contrast, mobile users must periodically signal their current location to the network, even if they don't have an active call. Mobility today also requires that the user's home network share the user's service profile with other networks. This can raise both security concerns for the network and privacy concerns for the user.

The ongoing transition from analog to digital cellular has not made a big impact on the traffic mix yet, but the next generation digital cellular systems, such as Wireless Asynchronous Transfer Mode (WATM) [IEE], may have a negative impact for several reasons. First, the cell-size will be much smaller. Current cellular systems often use cell-sizes of several kilometers. Next generation digital cellular systems will probably be "micro-cellular" with cell-sizes of less than 100 meters. This means that good models for user-mobility patterns become even more critical. The second reason is the bandwidth per subscriber. Current analog cellular systems (AMPS) allocate 30 KHz to each active subscriber, which is less than 1/50th of each cell's total available bandwidth. In contrast, WATM systems may have a bandwidth per cell of 20-30 Mb/s. In addition, one subscriber could use most, or all, of that cell's bandwidth. Small user groups tend to complicate statistical modeling, and hence traffic engineering. The final reason is that next generation digital cellular systems will also support multimedia traffic, such as MPEG video. As noted above, multiple media complicate traffic engineering.

Digital cellular companies are paying more attention to network security. In general, wireless channels are easier to attack than wired ones. Eavesdropping is virtually undetectable. Cryptographic techniques can provide privacy, but they usually entail a cost and performance penalty. In conjunction with mobility, unauthorized access is more difficult to detect and punish. So wireless system designers should include security in the earliest phases of the network design. The AMPS cellular system is one example where that was not done. (To be fair, AMPS was designed in the late 1970's and early 1980's before the combination of Microsoft's Windows and Intel's microprocessors brought cheap computing power to the masses.) AMPS uses clear-text signaling channels. Each handset has a Mobile Identification Number (MIN) and an Electronic Serial Number (ESN). For location management purposes, that data must be periodically broadcast to the cellular base-stations. As such, it can be easily stolen and -programmed into a "clone". That "clone" can then impersonate the real user, and hence steal service. The cellular companies now have numerous successful techniques for combating cloning. The simplest is PIN numbers that the subscribers enter while dialing a new call. This makes it harder to steal service, but not impossible. The need to use the PIN numbers also often annoys customers. Hence, other techniques have been introduced. A recent one is RF "fingerprinting". Each AMPS handset has a fairly unique RF signature, and the network will only accept calls if the ESN, MIN and RF signatures of the equipment match their stored values. Although this technique reduces cellular cloning [Bul], it is still not foolproof. So, a layered security-approach that uses PIN numbers, RF fingerprints, calling-pattern, location checking, and other similar information is still best.

Another security problem is that AMPS sends the user's speech as clear-text. This flaw has embarrassed several prominent individuals. Although cellular scanners are now illegal in this country, they are still widely available, as is information on converting legally available equipment to that purpose. [War] The AMPS fraud problem has been estimated at $700 million dollars per year in the U.S., which is about 3% of total industry revenue. [Bul] Digital systems which use the Cellular Authentication and Voice Encryption (CAVE)/IS-41C algorithm are more secure. Recent publications, however, have demonstrated weaknesses in both the voice privacy algorithm [Schw] and the authentication protocol[Pate]. Although it might appear that the answer for the AMPS security problem is to retrofit CAVE into every handset, that is not economically feasible. There are approximately 50 million of the AMPS handsets already in service. Security designs for public-network wireless systems should include a low-cost upgrade path.

Customer-Owned Equipment and Intelligent Content. In the older infrastructure, the regulated monopolies attempted to severely limit the types of equipment a customer could use on their network. They also centralized switching in their wholly owned network equipment. Those approaches are no longer feasible, or possibly even desirable, in the competitive environment. The current Internet philosophy is a prime

example of that shift. [Ste] That philosophy has the network flow control residing in the TCP layer in the end-systems, which then further complicates traffic engineering, since traffic peaks become less predictable.

As switching and other decisions migrate to customer-owned equipment, executable content becomes an issue. Readers are probably familiar with this problem in the context of Java applets or Word-macro viruses. The network infrastructure can also have similar problems. Proposed active networks may increase those problems. [Ten] In active networks, routers and switches perform customized computations on individual user data-flows. In addition, individual users can download and modify the programs that control those computations. As discussed below, the SS7 signaling network may eventually use active network concepts as it evolves into the so-called Advanced Intelligent Networks (AINs).

Vulnerabilities of the Physical Plant. Opening the network to greater competition may make general network services less vulnerable than the current infrastructure to physical attack if customers are aware of the vulnerabilities and diversify their telecommunications services across vendors. The economies of fiber optics have led to migration of traffic off diverse media (analog radio, digital radio, coax cable) and onto a few fiber optic routes. This minimizes right-of-way and maintenance costs while exploiting the ability to expand capacity by factors of ten by either upgrading electronics or putting several different colors of light down the fiber (Wavelength Division Multiplexing). These economies have subrogated any concerns for network-wide robustness. This was not true in the earlier integrated AT&T network (the "metro junction" plan). A few cuts of well-selected fibers would partition large parts of the nationwide telecommunications network. The same would be true regarding attacks on telephone equipment offices. Addition of a diversity of vendors, especially locally, may eventually enhance the connectivity of the long-distance telecommunications network. And the new diversity of local calling media (cellular, satellite, fiber) and diversity of protocols (circuit-switched, Internet phone) should make local phone service more robust.

The physical plant (support systems, physical equipment, etc.) is quite exposed to physical attack. In addition, there is no rigorous security culture within emerging telephony that reflects the growing importance that telecommunications in playing in the U.S. economy. By rigorous security culture we mean a culture that balances the management of the security of telecommunications with economic and service performance considerations. While pre-divestiture telephone companies had a strong security culture, the move toward more open systems, cost-based market forces, and extremely short time to market has eroded this culture. Surety is often at odds with minimized short-term cost and time to market.

Software Vulnerabilities. Over the last 30 years, telephony has been moving toward total software control of the installed switching and transmission equipment (e.g., operations, administrative, and maintenance, OAM, systems such as provisioning, monitoring, maintenance, billing; signaling control of circuit/packet routing). This change improves service performance and reduces costs, and increasing market pressures will ensure this will continue, with new software releases delivered in shorter development times. Unfortunately, the ability is lacking to ensure the surety of such an integrated software control system in the open telecommunications environment. The Year 2000 may well provide an example of the vulnerability of telecommunications to a pervasive software reliability flaw [You].

The open network of tomorrow will likely be less resistant to cyber-attacks than the previous closed system. Each telephony network, especially the regional telephone systems resulting from the breakup of the old AT&T, has on the order of one hundred major legacy operations, administrative and maintenance (OAM) software systems that control large portions of the network. Such systems were never designed to provide safe, secure access to the telephone network by outsiders. In fact, each of these systems is considered a black art in itself, and few telephony experts (some would say none), understand how such systems interact under all abnormal conditions. They do not exploit such principles as functional isolation to manage the creation of unintended vulnerabilities in the OAM architecture. The ability of an unauthorized user to discover a previously unconsidered system vulnerability is substantial. Unfortunately, software reliability and security engineering is in its nascent phases. There are strong theoretical reasons why it may remain that way for some time to come, so there may not be a technology fix for cyber-attacks on either new or legacy systems.

Changes to the Signaling Network. When we discuss the distributed network of the future, we need to describe those elements that will be distributed: transport, switching, and control. For transport and switching, distributed capabilities often mean a more robust network. A single network element failure may

lead to less, or perhaps no, service outage. If the network is properly designed, alternative routings and media are provisioned, and a coordination system for managing failures is in place and working. An open issue is whether the market will motivate a consortium of providers to coordinate the development of such a vendor-diverse system across their various networks – increasing network robustness at lower cost – or whether each carrier will perceive account control as too important. A potentially inefficient solution could result if each service provider develops his own vendor-diverse network.

Distributed control is another matter. It is usually difficult to guarantee the surety of coupled control systems. A number of well-publicized telecommunications failures have resulted from out-of-control, coupled distributed control systems (for example, see [Kuh]). In part, this results because it is difficult to get information on which to base a proper surety management system:

- Such systems are difficult to conceptualize or analytically model because they are highly complex, nonlinear in their behavior, and have many different inputs, outputs and processing functions;
- Performing a baseline characterization of them is difficult since many parts of the system may be changing at any given time;
- They are large enough so that organizational barriers get in the way of characterizing and understanding the entire coupled system.

Even if an accurate characterization of the coupled control system could be developed, there may be little hope of managing a real-time control system. If the adversary has access to perturb it in clever ways, the complicated nature and openness of the system may make it impossible to guarantee its invulnerability to attack in advance.

One goal is to have a single network for all multimedia traffic. The reality is that the Internet and phone networks will likely use separate backbone signaling networks for several more years. Both types of signaling have their own surety concerns. The old telephone infrastructure used in-band signaling which could be exploited with many tools. Out-of-band signaling, over closed Signaling System 7 (SS7) networks [Mod], eliminated that vulnerability during the 1980's. The 1990's, however, brought new problems for the infrastructure's signaling network. The first one was the return of in-band signaling: TCP/IP uses in-band signaling, since the control information rides in the same IP packets as the user data does. [Com] The classic exploitation of that is IP source-address spoofing, which became a substantial Internet nuisance. Authenticated TCP/IP headers could stop these in-band signaling problems, in theory. However, current authentication and key management techniques do not scale well to the planned Gb/s throughputs of the Internet backbone routers. The authentication problem is solvable at the local-access level though, if each ISP does key management for only its subscribers. In addition, the access speeds are much lower, so authentication may be technically feasible. Similarly, router-based source address controls can solve the IP spoofing problem to a large extent. This strategy has been applied in many of the newer high-speed network infrastructures, including some of the cable-modem infrastructure. Given the distributed, transnational nature of the Internet, however, not all ISPs are trustworthy. In-band signaling abuses may not have an ideal technical solution. As such, societal solutions must also be used. One example is the so-called "Internet Death Penalty", wherein other ISPs stop carrying traffic from an ISP that won't police its customers' behavior.

In the classical infrastructure, the signaling network of the telephone was "closed" in that only a few large monopolies and government-owed entities had access. In addition, employees from those entities wrote the software for the Service Logic Programs (SLPs) that ran on the SS7 network's Service Control Points (SCPs). Furthermore, the lack of time-to-market concerns allowed extensive testing of the programs. Restructuring has changed this closed system for several reasons [Cas]. The first reason is that third-party connections to the SS7 network are mandated by law in the restructured "Advanced Intelligent Network" (AIN). Those third parties may include customers as well as other service providers. The customer interfaces to their service parameters may be via the Internet. As such, the SS7 network may gradually acquire all of the Internet's security problems. One example is one plumber's call-forwarding attack that stole business from rival plumbers. [NYT] The second reason is time-to-market concerns that dictate rapid service deployment in a competitive marketplace. Rapid deployment, though, often means lower-quality software. Consequently, improved software methodologies are a critical research area. The final reason is that most SS7 traffic is sent as clear-text. That traffic includes sensitive data such as calling-card PINs and

credit card numbers. So, as the SS7 network becomes more open, data privacy may require cryptographic techniques.

Another future direction of the telephone network is to move away from channeled circuit-switched network towards TCP/IP or ATM networks lying on high-capacity transport. These networks operate on packets. If traffic routing control for these packet networks is provided to each competing vendor, then a malicious consortia of vendors could adversely affect traffic by manipulating the routing and creating congestion or failure conditions in the network for competitors, or perhaps for the entire network itself. This has its equivalent in market manipulation where the actions of a cartel can impose inefficiencies for nonmembers. The challenge is to impose real-time controls that enhance stability and fair operation of the network bandwidth marketplace while being minimally intrusive.

Remote Maintenance and Outsourcing. Competitive service providers are under intense pressure to reduce the size of their operations and maintenance staff. These reductions usually take two forms -- remote maintenance and outsourcing. Remote maintenance allows a centralized help desk to perform troubleshooting and some repairs and obviously lowers costs. It can also improve surety. Continuous monitoring can often uncover transient and systemic problems before they affect service. On the other hand, remote monitoring via the Internet or dial-in modems also introduces all of the standard security problems of computer networks into the telecommunications infrastructure. The executable content problem may also appear. For example, applet-based remote-monitoring probes might download executable content into the ISP backbone routers.

Outsourcing replaces an in-house maintenance staff with either on-site contractors or a service contract. The service contracts are typically with either equipment vendors or firms that specializes in network maintenance. These service contracts, which often mandate remote access to the customer's network, entail the tradeoffs mentioned above. However there are additional surety tradeoffs. The contractors are usually quite knowledgeable about particular hardware/software platforms. Also, their multi-company focus often exposes them to a wider range of solutions than in-house staff. These benefits, however, are often balanced by their lack of detailed knowledge of, and long-term experience with, the quirks of a particular customer's network. Finally, maintenance personnel often have access to sensitive company data as it passes through the company's infrastructure. One long-distance company has already had a problem with an employee acquiring calling-card numbers off of their SS7 network. It is unclear if outsourcing will increase that insider-fraud problem, or not. Outsourcing certainly introduces a number of challenges to surety, including but not limited to, incompatibilities in personnel security policies, a lack of employee loyalty, surety culture, and group cohesion, an ideal location for the placement of agents, and the potential ability for a smaller number of attackers to impact a larger number of systems.

Human Factors Concerns. Introduction of technology used to be relatively slow and measured, and training was thorough. A single vendor made most telephone equipment, and the carriers and equipment providers had a large and experienced staff to check and recheck. The telephony future looks like more vendors demanding more services from lean-and-mean data-freight providers, in a market with greater vendor and service churn and greater demand for quick turnaround in provisioning services. This will lead to more people reconfiguring, either physically or logically, various network elements. To provide for the new services that opening the network is expected to stimulate, there will be increasingly quick introduction of new hardware and software. The combination of these effects will lead to more mistakes being made during operations.

There is no single vendor who is going to take the lead, as the old Western Electric had, in improving the human factors of such equipment. It was difficult enough to translate the economic costs of poor operations through to effective equipment design when there was one manufacturer, one service provider, and appropriate incentives. With a multitude of suppliers and carriers, it is overly optimistic to assume the market will make things right, at least in the near term.

The Risks of Reliance on Redundancy. Not only is redundancy not a panacea, but excessive reliance on it can ultimately make a system less reliable, less safe, less secure--less sure. To have redundancy implemented correctly, the designer must include the initial design as well as the construction processes and operational procedures in his analyses--the entire system as designed, as built, and as operated.

People often count on redundancy without examining the system in sufficient detail to identify all the potential common mode and common cause failures. Seeing what looks like redundancy on a more superficial level can make management complacent. Digging deeper, however, can often reveal is a hidden shared support requirement among "redundant" systems or components or unrecognized interconnection between them. And, the more redundant pathways that are built into a system, the more potential exists that at least some of them share interconnections or support.

This is not just academic. Let two examples from transportation suffice. First is the crash of the just barely controllable United Airlines DC-10 at Sioux City, Iowa on July 19, 1989. The DC-10 had three "redundant" hydraulic systems, or so the designers and airplane certifiers believed. All three, however, shared a particular location, and a single malfunction--a fragment from a failed engine rotor--was the common cause for all three to fail concurrently. The systems were not redundant with respect to this fault. [NTS]

Second, the Morton-Thiokol designers and NASA engineers thought that the two O-rings on the solid rocket booster used for Challenger were "redundant," in all but one possible scenario. That scenario was deemed to be extremely unlikely. For all other situations, their models predicted that either of the O-rings would individually provide the necessary seal. Unfortunately as we now know, the models were wrong, and there was another scenario in which the redundancy failed, an additional common cause for double seal failure.[Vau]

While redundancy may provide coverage for certain classes of faults, redundancy also increases component failure rates. Redundant communications paths increase the opportunities for leakage and other attacks while increasing the cost of protection of these now separate and different components. The elimination of common-mode failures means that we need to do redundant engineering, maintenance, and acquisition, and that we need to produce smaller numbers of more different components. Each of these leads to inefficiencies from a cost perspective.

Security vs. Privacy. Network surety has two competing aspects. The first one is society's desire that individuals or groups not use public resources for nefarious purposes. The other aspect is a fundamental right to privacy for an individual. These privacy concerns can revolve around personal data, such as health, purchasing and entertainment records, in addition to fundamental political rights, such as free speech. Hence, the appropriate political compromise between security and privacy is still unclear. Strong cryptography is one solution to these privacy concerns. [Pla] Encryption allows individuals to communicate with reduced overall risk of eavesdropping. PGP remailers allow anonymous web-surfing and USENET posting. With the growth of the Internet culture, those cryptographic techniques are now widely available from multiple vendors in multiple countries. These same cryptographic techniques, however, also abet organized crime and terrorist organizations. Some national governments have instituted bans on private cryptography. Others have proposed either key escrow systems or weak cryptographic protocols for their nationals. While those proposals reduce individual privacy, they can also place a country's industry at a competitive disadvantage. Enforcement is also difficult, given the global nature of the infrastructure. Finally, weak cryptographic protocols may be technically dubious. Recent Internet experiments have shown that even 56-bit DES is vulnerable to concerted attacks with widely available hardware and software.

If access to physical plant such as offices is provided and there is a greater availability of high-bit-rate services by vendors, then there will be greater opportunity for a malicious vendor to access high-bit-rate lines without authorization and to compromise the integrity, availability, and confidentiality of hundreds of voice lines or voice-grade private lines.

A more open and distributed future for telecommunications is both good and bad from a reliability and security standpoint. There may be more routing options, especially for the vanilla data-freight services, although it will still be the responsibility of the customer or some third-party provider to design their vendor-diverse network robustly. But many of the major failures of the last few years have been either software failures, supporting infrastructure failures or subtle system failures. Many signs point to the telecommunications infrastructure becoming more vulnerable as restructuring and evolving new technologies open the network to rapid, uncoordinated change.

Information Warfare. As mentioned above when discussing the external threats to the U.S. telecommunication infrastructure from terrorists, criminals, and others, one of the threats they pose is "Information Warfare "(IW). This term has a distinct military flavor due to the word "warfare." One

military definition of IW is: "Actions taken to preserve the integrity of one's own information systems from exploitation, corruption, or destruction while at the same time exploiting, corrupting, or destroying an adversary's information systems." Use of this term by the military has been replaced with "Information Dominance," which represents non-wartime protection of U.S. assets and preparation for conflict with potential enemies. "Information Warfare" is then reserved for active operations against enemies.

IW is a fairly broad term which has been interpreted to represent a number of activities which directly affect the integrity, use, and flow of information or which affect information in order to produce some secondary outcome. Examples include the following:

- hacking
- diversion
- theft of services
- loss of information
- bank fraud
- cyber crime
- corporate espionage
- blackmail
- unauthorized use
- propaganda
- terrorist activity
- economic instability
- military attack

In an earlier section we discussed this threat in general.The range of possible IW attacks can also be explored by examining the motivation of the attackers in more detail. For the system designer and operator another characterization of IW attacks is by how an attack is implemented. The many methods of attack can be divided into physical and non-physical (information) attacks. These attacks may result in both physical and non-physical impacts. Examples are shown in the following table.

		Changes caused by attack	
		Information/non-physical	***Physical***
Attack	***Information/ non-physical***	1. Unauthorized remote access to telephone switching system configuration allows denial of service or theft of service 2. Computer virus is activated which could stop, modify or purge record of financial transactions	1. Telephone denial of service outage brings down air traffic control communications which slows or stops flight departures and landings 2. Access to a satellite control system jettisons stabilization fuel which reduced life of satellite or causes immediate loss of use.
	Physical	1. Damaging communication cable or microwave tower prevents communication over the link. 2. Damaging of one or more communication links causes remaining links and nodes to be jammed with rerouted transmissions, slowing and stopping certain communications	1. Bombs cause damage to switching equipment in strategic locations which is difficult and expensive to replace and causes denial of service 2. A telecommunications switching center stops working for no apparent reason. A truck just drove by and emitted a high-power microwave pulse, damaging sensitive electronics.

Terrorist IW activities can be grouped into at least three general categories. They can be used to disseminate propaganda, to raise funds, and to attack the infrastructure itself. When telecommunications is the infrastructure under consideration, the first two of the three also indirectly affect the infrastructure in that the unauthorized use of telecommunications channels to either spread falsehoods or to steal or extort money would lessen the confidence of the public in the integrity of the telecommunications network. While confidence can be restored, it is not instantaneous nor is restoring it without cost, including the cost of lost revenue due to public avoidance.

Unless an attack is physical, (e.g., a bomb), the attacker must get access to the system (e.g., access a telephone). Hence, a successful attack must first penetrate and then disrupt the normal operation of the telecommunications network. Rathmell has written that

> Terrorist activities in cyberspace may be considered as part of a new kind of war: software warfare. When InfoWarriors plan to hack or penetrate particular networks, their goal is to modify software and, consequently, its proper functions. Conversely, the system managers of the targeted information systems have to make sure that software is protected and running properly. Other forms of Information Warfare, such as Command and Control warfare, Information Infrastructure Warfare, or economic information warfare are therefore dependent on the outcome of this competition to control the software of information systems.[Rat a]

This software war--this war for control of the telecommunication system—may involve access to an authorized user's password. This can be done, for example, physically by theft from or extortion of an insider, treacherously by an insider willingly cooperating, or electronically by a software "sniffer". After penetration, if the attacker wants the capability of the system manager or its equivalent, they might introduce a Trojan Horse. A Trojan Horse is a program that has undocumented side effects. The attacker might also create a software bomb to destroy logic or files at a later point in time. Another possibility is viruses (by the end of 1996, about 8000 virus strains were known). They could be used to spread panic or could reproduce in such numbers as to cause gridlock. [Rat a]

The concept of information warfare is not new. After all, the ancient Chinese strategist Sun Tzu wrote about confusion, concealment, and deception. As Lawrence Freedman of King's College has wryly observed, "Large explosions are a traditional and effective form of information warfare when directed at key transmission or storage facilities."[Fre]

What is new is the evolving ubiquity of the system that can be attacked, and the increasing reliance of government and business on the timely flow of accurate information from that system. The telecommunications system has so many interconnections and is handling so much traffic that a perturbation of it may have great effect, both in rapidity and in breadth of activity, on our lives. Again, Freedman notes that "As with so much else modern technology appears to have moved us well beyond the inefficiencies of yesteryear but at the price of magnifying the consequences of even a single malfunction."[Fre] Malfunctions happen everyday, but they can also result from an attack, and attacks can be designed to make the negative effects of the malfunction much greater than those of a single point failure.

We should not make the success of such an attack seem easier than it would probably be. The telecommunications system should not be assumed to be an easy target. "Common sense has led most users of modern information technology to take precautions." [Fre] For example, from the software side of things, we make backups and control access to our systems. We search for viruses. Yes, penetrations have occurred and still occur, but detected attacks have been primarily attributed to hackers, not "cyber-mercenaries." According to Rathmell, those who warn of IW

> foresee concerted, strategic attacks on the [telecommunications infrastructure] by foreign states or organized terrorist groups. They envisage an enemy able to crash the telecommunications system, disrupt air and rail traffic control systems, subvert the financial system and undermine the power distribution network. All of these activities are possible, and the tools and techniques for doing so are readily available on the Internet. Carrying out such an attack would however require a sophisticated intelligence study of the targets, something which only a few states and organizations could contemplate. [Rat b]

To be successful, the attacker needs a competent attack capability. They are computer literate people who are "...highly skilled and trained products of government agencies or corporate intelligence branches working on the open market."[Rat a] In addition, the IW attacker requires an extensive knowledge of computer networking and must do detailed intelligence work of the network to be attacked.

The planning of such attacks requires resources, and thus, it is more likely that a larger organization would be the sponsor of IW. Would people do it? Medd has written

> Terrorists of the future will likely intensify their targeting of the world's financial resources that are being transferred daily over the information highway. By using and attacking these financial conduits, terrorists will be able to transfer funds from their sponsors as well as to tap illegally into the most vulnerable legitimate transfers of others. They could conceivably cripple the financial markets either by direct attacks or by instilling doubt in the electronic financial world.[Med]

The goal of IW attackers "...would be the disruption or destruction of information infrastructures including basic services such as power supply, police databases, social security transfers, medical networks, transportation signals, money transfers and telephone switching systems."[Rat a]

Terrorist organizations and rogue states may well prefer to attack in cyberspace. The risks to the attacker are reduced, they can attack from a distance, and it is more difficult to prove causality. Freedman points out, however, that, if an enemy is not sure that he has the computer expertise at his disposal, as well as the right analyses, he ".might well be unsure about relying on clever and subtle forms of electronic warfare to disable a critical facility, especially when something cruder, simpler and probably more violent will do. Why become a hacker when it is as easy to be a bomber?" [Fre] There is the difference, noted above, that a bomber must physically transport the bomb to the site whereas an information warrior could do so with only bits being physically transported.

MITIGATION STRATEGIES

Issues have been raised above that indicate that surety issues may challenge the future of telecommunications. In this section, we introduce several mitigation strategies, strategies that we have found useful in many different applications to surety issues that can lessen the likelihood or severity of "accidental experiments." These strategies are not mutually exclusive; in fact, the best approach that we have found to achieving a sure system is to incorporate a number of them in the design.

Response to Normal Accidents

The first principle for designing systems is to make them reliable by using reliable parts. But more must be considered in the design. Birman writes, "Even if every component of a system were extremely dependable, the story would not end there. Merely interconnecting reliable computers and bug-free programs does not yield a robust distributed system. Instead it produces a network that works well under most conditions....Some additional form of protection is therefore needed."[Bir] This is particularly true when the possibility exists for intelligent attack. What programmers have developed is the same as what communications network designers need to continue to develop—fault-tolerant systems that can rapidly re-configure themselves to bypass failed elements, whatever the failure cause. In the case of intelligent attack, the volume of the network attacked might very well be much larger and detecting the attack could well be more difficult, but the principles for continued system success are the same. The question then is "How do we do this?"

Recall that Perrow considered the susceptibility of a system to have an accident a function of two parameters: complexity and tight coupling.[Per] Kuhn has observed that the telecommunications system, while it is quite complex, is actually relatively loosely coupled: "In most system, a trade-off can be made between simplicity of interactions and looseness of coupling. We can consider the PSTN [the public switched telephone network] a loosely coupled system because it can dynamically reroute calls along many paths. However, it achieves this loose coupling at the cost of some complex interactions between components." He goes on to write, "For a communications system, coupling is probably the more important of the two properties [the other is complexity] in determining its capacity to tolerate failures. It is directly related to the system's primary function: maintaining connections between points."[Kuh]

We assert that a means of designing loose coupling into a system, whether complex or simple, is to incorporate the concept of the 3 *I*s (or 3 *D*s) into the design. At Sandia, we do this by the creation of a "safety theme" for a system that addresses the 3 *I*s and demonstrate how that theme is achieved by the specific design. A possible usage of the concept in telecommunications is to design in *isolation* so that

faults do not propagate throughout the system and result in a scenario like the northeast U.S. electric power blackout of 1965. For authentication of the communication, make the key *incompatible* with other signals occurring in the system. If faults occur and start to propagate, one can design that portions of the system become *inoperable*. It may well be better to lose part of a system for a short time then all of a system for a long time, and some customers may be willing, for a price, to be separable. There are obvious analogies here to the electric power grid and the concept of selective blackouts. For security concerns, design the system that an unauthorized attempt at access is *delayed* as long as possible, including detection and ultimate defeat in this scope. Provide *discrimination* between legitimate and illegitimate signals. Again, if access is gained, *disabling* part of the system may be the preferred alternative to contaminating the whole.

Consequence-Based Analysis

Portions of the nation's infrastructure are vulnerable to a wide range of potential threats, ranging from aging and degradation to physical attacks to malicious intrusion of software systems, whether computers or telecommunications switches. While aging and degradation threats are real and can be significant, we shall focus in this section on intelligent threats and a strategy to address them called "consequence-based analysis." An exact definition of such threats to infrastructures, complete with probabilities of occurrence, would enable the U.S. to efficiently design protection and countermeasures. Such a definition would also support the identification of critical nodes and allocation of resources to the most important areas. Unfortunately, the definition of such a threat has been elusive. Few specifics are known although we have these:

1. The number of computer security incidents has substantially increased
2. Cyber attacks are hard to detect and trace and are plausibly deniable.
3. There are indications of wide spread "listening."
4. Terrorists, organized crime, and foreign governments have access to sophisticated technology.
5. Cyber threats to an organization often arise from within the organization.
6. Analysis of the new threat may require new approaches by threat analysts.

It is unlikely that this nation, or any other nation or international body, will have a threat definition in the foreseeable future that is sufficiently detailed to be the sole basis for infrastructure protection decisions. The problem of threat definition is compounded by the fact that the nature of the threat is ever changing. Changes can be caused by:

1. Advances in threat technology (affects an adversary's capability)
2. Changes in the local, national or international political climate -- some of which may be unknown and unknowable to the analyst (effects an adversary's motivation)
3. Advances in deployed protection technology (as one "hole" is closed, an adversary may seek out another point in the system to exploit), and
4. Advances in the perceived level of system protection (if a system is perceived to be too tough to attack, an adversary may seek to attack another system that is perceived to be an "easier target").

These factors combine to demonstrate that it may never be possible to have an adequate and workable threat definition. Even if such a definition could be created, it would be at best difficult to demonstrate its completeness, and the task of keeping it up to date would be formidable.

The protection of critical infrastructures is vital to our nation's ability to effectively respond to both natural disasters and hostile acts. However, the lack of a detailed threat definition has historically limited the degree to which private industry is willing to invest in infrastructure protection. Protection cannot be effectively achieved if industry is not involved, but industry requires, and has thus far not received, a clear business case for investing in protection against high-grade threats. Furthermore, the proper role of government in protecting privately held assets is being debated. Without some clear rationale for a decision as to where responsibility lies for protection, the question is difficult to resolve.

Sandia has developed a consequence-based approach to the identification of critical nodes that has found acceptance with some businesses. For example, this approach is currently used by insurance companies at a high level and has been used as the basis for the assessment of nuclear power plant safety. The approach incorporates the probabilistic risk assessment methodologies that Sandia has developed for and applied to the problem of designing safe, secure nuclear reactor facilities. The consequence-based approach provides an identification of critical nodes from a system perspective. The definition of "critical" is based on an understanding of the consequences of a failure of a system or component of the system. Not all vulnerabilities are critical. When development of this method began, there was no clear definition of the threat, but it was recognized that certain consequences (e.g., the release of radioactivity to the atmosphere) were unacceptable to the public. Avoidance of these consequences was accomplished by starting from a systems design approach, understanding the system operation, identifying failure modes, assessing the risk of system failure, identifying critical nodes, and designing safety features. This process allows identification of risk so that the costs of protection are viewed in the context of acceptable risk. Similar approaches appear to be beneficial for enhancing the security of telecommunications.

Vulnerability Assessment

To prepare for a malicious attack, one may perform a vulnerability assessment, which discovers specific technical vulnerabilities in a system and the components of which it is comprised. To do this well, one must understand thoroughly the spectrum of technologies involved, have insight into the community which has the skills to cause damage, and understand the groups/nations that might which to utilize these skills [Rat a]. Ideally, such assessments are performed by an independent entity, since the judgement and understanding of the system owners may lead to assessments that are colored by the same culture that influenced the system designers. Internal self-assessments generally lack the "out of the box" thinking and deep maliciousness that could characterize an external entity determined to do harm.

The results of vulnerability assessments must be kept tightly held since the mitigation of specific vulnerabilities may not be completed rapidly. Additionally, the system owner may feel that the probability of attack for certain vulnerabilities is low enough and cost of mitigation high enough that leaving the vulnerability in place is an acceptable risk. Widespread knowledge of the vulnerability would adversely affect the probability of attack. The background and proven discretion of the individuals performing the assessment must be understood.

The spreading telecommunication network has not had system-wide security as a major design requirement [Rat b]. In general, vulnerability assessments such as described above have not been performed at a level that would capture critical systems-level vulnerabilities. Identifying the vulnerabilities of the entire system is a common good, but not necessarily good for a particular company vs. another company. Hence, this may be an area where cooperative action involving business and governments may be warranted.

Hardening And Survivability Of Management Layers

As in all mitigation strategies for the U.S. infrastructure, protecting the management layers should involve inclusion of surety both at a component and a systems level. Concepts discussed here are included to provide examples. Independent improvements intended to secure information, but added outside of a systems perspective, may be ineffective at improving surety of the infrastructure. Improvements must be considered at a detailed systems level to avoid overlooking vulnerabilities, or introducing new ones.

Exact configuration of management layers may depend on particular location, companies involved, and technologies selected, and may also depend on the time a particular provider connects to the infrastructure and which regulations are in effect. To discuss mitigation strategies here, the management layer is divided into several conceptual pieces. The first piece is the physical connections, links, and equipment in the management network which carry management information. Second, independent command and control functions exist in various systems and equipment which are accessed by a limited number of authorized owners or customers. These systems control equipment and network operation in a way which normal providers need not access or control (such as fundamental routing information, computer codes for network operation, etc.). The third is a network-based command and control network to which all providers attach to route information.

Hardening and survivability of the first component, the physical components which carry management information, requires protection from unauthorized access and modification. If an adversary is

able to deny, alter, or even access this information, it could be used to adversely affect the communication network, the providers, or customers. Solutions here may include adding authentication information to the management data so attempted alterations would be incompatible with the system. Adding encryption would make it very difficult to eavesdrop or alter management data. Monitoring the links for state of health and providing redundancies may also improve the physical links. Equipment selected for use in a network may have inherent vulnerabilities. These may be identified by an investigation and certification review of the equipment using a systems approach.

Hardening the second component, functionality of independent system command and control, is often reliant on surety concepts used for remote and local access to computer systems. These systems are typically more secure because they allow limited access by a small set of authorized persons. Often passwords are used to control accounts. Is access available remotely over a network, the Internet, a telephone line or is only local access available? Are these links protected? Are the networks secure? Vulnerabilities of these systems include theft of passwords, password guessing, eavesdropping on links, link hijacking, and denial of service attacks. Again, the ability of the system to discriminate between authentic and unauthentic messages is paramount.

The third component has the problem of access by the widest range of individuals and therefore is vulnerable because it is more likely for one of these people to be an adversary or make a mistake due to incompetence. Limiting functionality at this level is a useful tool to limit possible damage to the communication network, yet it will cause more overhead and creation of an elite command and control segment. As decisions are made about this tradeoff, other methods of improving surety may be considered. Including systems designed with multi-level access, accountability (audit trails), and automated monitors for improper use may improve the inherent surety of such a network.

All of these three conceptual segments could benefit from addition of physical security measures. Physical security can be implemented in many ways. Access to particular links and nodes can be disabled except for those specifically authorized. Facilities, equipment, and links are located away from streets and parking lots which may be bombed. Rooms holding important equipment are locked and entry only granted to those who need access. Basic steps used to protect from lightning and radio interference may be followed to avoid damage from electromagnetic pulse weapons. Personnel are investigated to help eliminate potential adversaries from holding trusted positions. Important concepts to consider in physical security are detection, delay, and response. A simple lock may protect access to a room; however, an alarm system helps enhance this security. It detects unauthorized entry into that room. Delay mechanisms can be used to prevent further access--or even retreat by the intruder--until some response can be made. Response may be sending security personnel or simply video-recording the event and intruder. These measures help deter unauthorized access.

Indications and Warning System

The Defense Science Board postulates that it is possible for an adversary to mount a structured attack against infrastructures while disguising the attack as a series of apparently unstructured, random events. This scenario is plausible because there is no capability within the U.S. to coordinate among the infrastructures an Indications and Warning (I & W) of physical or computer-based attacks. Providing I &W of attacks in a timely fashion may be the most difficult technical and organizational challenge facing those charged with protecting the U.S. infrastructure.

Traditional strategic and tactical warning systems may not be useful in providing Indications and Warnings of attacks against the U.S. infrastructure. The infrastructure threat is vaguely defined in parts of industry and the ability to determine the extent of an attack on the U.S. does not currently exist. The technical and organizational challenges are formidable. For example, the information necessary to provide the indications of an indirect (subtle or below threshold) or a direct attack would have to come from private sector organizations who own the various infrastructure elements. Real-time monitoring of system operating parameters is common in most of the infrastructures but the analysis and response of abnormal events is focused on maintenance and restoring operation of the system. For an I&W system, the response must focus on restoring operation of the system as well as providing warnings to the infrastructure under attack as well as the other infrastructures which might be impacted and identifying the adversary for criminal prosecution. In addition, certain infrastructures such as the telecommunications industry must contend daily with numerous attempts to illegally enter their systems using the computer as a criminal tool. This burden of providing security to these infrastructures whose needs differ greatly emphasizes the

importance of an automated system which recognizes, records and stores selected events for analysis. The organizational challenge lies in creating a structure where industry would agree to the monitoring and reporting of this data among competitors, law enforcement officials, and other elements of the infrastructure. Furthermore, the nation needs the capability to monitor the interdependencies among the infrastructure elements and use this information to improve the ability to assess the nature and severity of an attack. Private industry and government must cooperate to achieve the objective of providing adequate response time to assess an attack situation and respond accordingly.

The I&W problem is by its nature a very large distributed collection of technical and organizational issues. The approach should be to a) identify the major participants needed for the I&W concept, b) determine the technical problems to be addressed with priorities of importance and c) begin coordination of the organizations needed to implement the goals of the I&W concept.

The features desired in an Indications and Warning system include a) detection of precursors to an attack, b) awareness that an attack is in process, c) correlation of disparate attacks, and d) recommendations on how to recover from an attack and, perhaps, identification of perpetrators.

The activity of an I&W effort should be guided by a coordinating center which is the focal point of a large matrixed activity. The center would provide the coordination among the infrastructures and liaison with the law enforcement community. Analysis by the coordinating center would result in attack warnings provided to both industry and the law enforcement.

Other Mitigation Possibilities

Technology is advancing and revolutionizing telecommunications. At the same time, its advancement will also deliver new mitigation techniques, ones that we cannot foresee at this time. As an example, recent research by Rashari Roy at the Georgia Institute of Technology suggests that encoding messages into the background noise found in fiber-optic cables has the potential to increase security of information transmission. The transmitter and the receiver would both need to know the encryption protocol, but without it, no message could be detected within the noise.[Pri] Other examples of innovation shall undoubtedly occur in the future.

CONCLUSION

The challenges to the telecommunications industry and its stakeholders appear daunting. Above, we have briefly discussed how this industry is rapidly changing in the U.S. It is also doing so worldwide. Global infrastructure surety will be increasingly dependent upon the various national infrastructures. Indeed, the notion of a discrete national infrastructure is losing meaning. The telecommunications infrastructure will dramatically reconfigure to incorporate rapid technological advances, will probably grow orders of magnitude in terms of the numbers of competing providers, and must respond to an uncertain regulatory and legal framework. Surety is easiest to engineer into discrete, well-understood systems. Indeed, the exceptional reliability and acceptable security of the current infrastructure are a result of the systems engineering environment of the past. The future environment of multiple providers, multiple technologies, distributed control, broad access to hardware and software is fundamentally different. The solutions that will underlie the surety of the telecommunications infrastructure of the future will be shaped by this different environment, and may be expected to differ from the solutions of the past.

Current policy discussions tend to treat the infrastructure's surety as an expected product of market forces. Where there are demands for high reliability or high security, there will be suppliers—at a price. Indeed, end customers will have an unprecedented ability to chose to protect themselves by buying services that can function as a back-up, demanding services that support individual needs for security, and choosing proven performers as suppliers. However, pervasive confidence in the infrastructure, comparable to the expectations that are a result of the surety of today's infrastructure, is not guaranteed. The surety of the future global telecommunications infrastructure could well be significantly greater that that of today for both specialty needs and as a collective average. It could also be lower, if market forces do not result in surety requirements comparable to those met today. The ability of the marketplace to anticipate and address public expectations around the low probability of occurrence but high consequence events that could result from abnormal or malevolent environments remains to be seen. We are moving from an era where the

surety of the infrastructure was generally predictable and controlled to one in which there are profound uncertainties.

Generally, the private sector does not expect that the market will provide the level of security or resilience that would be required to mitigate a serious information warfare attack on the telecommunications network or other essential pieces of the infrastructure. The issue of private sector and public sector rights and responsibilities as relates to the surety of the telecommunications network remains an intriguing policy arena. Within the U.S., the government serves the role of both regulator and concerned customer. Essential governmental functions including continuity of government, emergency services, and military operations depend upon the surety of the telecommunications industry. Institutions such as NSTAC and the NCC have provided an effective forum for working these issues cooperatively. Their role and their effectiveness in the dramatically changed policy, business, and technical environment of the future remain to be defined. Recent initiatives external to government include the recent Manhattan Cyber Project, described in a press release as "a concentrated outreach initiative between industry, government, and academia to address the cyber threat impact on the National Information Infrastructure and competitiveness of corporate America." The recommendations of the President's Commission on Critical Infrastructure Protection [PCC], and the public and private sector response to those recommendations, are critical to anticipating the future public/private sector relationship as to the surety of the telecommunications piece of the infrastructure.

REFERENCES

[Act] "The Communications Act of 1934 as Amended by Congress 1996"

[Ber] N. Berstein, "High-Tech Sleuths Find Private Facts Online," *The New York Times*, September 15, 1997.

[Bir] K. Birman and R. van Renesse, "Software for Reliable Networks," *Scientific American*, May 1996.

[Bul] W. Bulkeley, "Stop Thief", *Wall Street Journal*, September 11, 1997.

[Cai] F. Cairncross, "Telecommunications," *The Economist Survey*, September 12-18, 1997.

[Cas] T. Casey, "The Advanced Intelligent Network - A Security Opportunity", 19th Systems Security Conference, Baltimore, MD, November 1996, pp. 221-232.

[Com] D. Comer, *Internetworking with TCP/IP, Volume 1, Principles, Protocols and Architecture, Third Edition*, Prentice Hall, New Jersey, 1995.

[Cra] R. Crandall and J. Ellig, *Economic Deregulation and Customer Choice: Lessons for the Electric Industry*, The Center for Market Processes, Fairfax, Virginia, 1996.

[D'An] P. D'Antonio, "Surety Principles Development and Integration for Nuclear Weapons," *Proceedings of the Second International High Consequence Operations Safety Symposium*," Sandia National Laboratories, July 29-31, 1997, Albuquerque, New Mexico, in preparation.

[Dav] J. Davis, et al., "Achieving High Assurance, High Consequence Systems with Model Integrated Computing," (to be published)

[Eco] *The Economist*, "Surviving the Telecoms Jungle," April 4-10, 1998.

[Fis] M. Fisher, "Technology: moldovascam.com," *The Atlantic Monthly*, September 1997.

[Fow] W Y. Fowlkes and C. M. Creveling, *Engineering Methods for Robust Product Design*, Addison Wesley, Reading, Massachusetts, 1995.

[Fre] Lawrence Freedman, "Information Warfare: Will Battle Ever Be Joined?," Lecture given at the Launch of the International Centre for Security Analysis, London, England, October 14, 1966.

[Ger] J. Gerth and J. Miller, "Funds for Terrroists Traced to Persian Gulf Businessmen," *The New York Times*, August 14, 1996.

[Giu] B. Giussani, "International Alphabet Soups Seek to Regulate Internet and E-Commerce," *The New York Times*, March 31, 1998.

[Har] Harvard Group on Risk Management Reform, *Reform of Risk Regulation: Achieving More Protection at Less Cost*, Center for Risk Analysis, Harvard School of Public Health, boston, Massachusetts, March 1995.

[IEE] "Wireless ATM", *IEEE Personal Communications*, August 1996.

[Kai] J. Kaiser, Chairman, Accident Investigation Committee, "Flight 965--Accident Investigation Summary," Airline Pilot's Association, August, 1997.

[Kal] A. Kaletsky, "Snakeoil, software, and Gates," *The London Times*, September 10, 1997.

[Kle] L. Kleinrock, *Queueing Systems*, Chapter 3, John Wiley, New York, 1975.
[Kuh] D. R. Kuhn, "Sources of Failure in the Public Switched Telephone Network", *IEEE Computer Magazine*, April 1997, pp. 31-36.
[Kun] D. M. Kunsman and D. D. Carlson, "Approach to Analyzing the Inadvertent Operation of a Standby System," *Proceedings of the Probabilistic Safety Assessment International Topical Meeting*, American Nuclear Society, Vol. 2, p. 1404-1409, Clearwater Beach, Florida, January 26-29, 1993.
[Laq] W. Laquer, "Postmodern Terrorism," *Foreign Affairs*, 75, 5, pp. 24-37, New York, September/October 1996
[Lel] Leland, et al, "On the Self-Similar Nature of Ethernet Traffic", *IEEE Transactions on Networking*, February, 1994, pp. 1-15.
[Mac] S. MacLeod, "The Paladin of Jihad," *TIME*, 147, 19, May 6, 1996
[Med] R. Medd and F. Goldstein, "International Terrorism on the Eve of a New Millennium," *Studies in Conflict and Terrorism*, 20, 3, 281-316, Washington D.C., Sept., 1997.
[Meh] A. Mehrotra, *GSM System Engineering*, Artech House, Boston, 1997.
[Mod] A. Modarressi and R. Skoog, "Signaling System No. 7: A Tutorial", *IEEE Communications Magazine*, July 1990, pp. 19-35.
[Neg] N. Negroponte, *Being Digital*, Vintage Books, 1996.
[NTS] National Transportation Safety Board Docket No. DCA89MA063 stored in NTSB Microfiche No. 39915A.
[NYT] *The New York Times*, "Plumber is Charged with Reaching Out Illegally for Customers", January 29, 1995.
[OTA] Report No. OTA-E-453, *Physical Vulnerability of Electronic Systems to Natural Disasters and Sabotage*, June 1990.
[PatC] M. E. Pate-Cornell, "Learning from the Piper Alpha Accident: A Postmortem Analysis of Technical and Organizational Factors," *Risk Analysis*, 13, 2, Plenum Publishing, April 1993, 215-223.
[Pate] S. Patel, "Weakness of the North American Wireless Authentication Protocol", *IEEE Personal Communications*, June 1997, pp. 40-44.
[PCC] *Critical Foundations: Protecting America's Infrastructures*, The Report of the President's Commission on Critical Infrastructure Protection, October 1997.
[Per] C. Perrow, *Normal Accidents--Living with High-Risk Technologies*, Basic Books, 1984.
[Phi] C. Phillips, Jr., *The Regulation of Public Utilities: Theory and Practice*, Third Edition, Public Utilities Reports, Washington, D.C., 1993, p. 766.
[Pla] C. Platt, "Plotting Away in Margaritaville", *Wired Magazine*, July 1997, pp.140-144, 176-179.
[Plu] D. W. Plummer and W. H. Greenwood, *A Primer on Unique Signal Stronglinks*, SAND 93-0951, Sandia National Laboratories, August 1993.
[Pri] M. Prigg, "Noise Keeps Money Safe," *The London Sunday Times*, March 1, 1998.
[Rap] T. Rappaport, *Wireless Communications*, IEEE Press, New York, 1996.
[Rat a] A. Rathmell, et al., "The IW Threat from Sub-State Groups: an Interdisciplinary Approach," *Third International Symposium on command and Control Research and Technology*, Institute for National Strategic Studies, National Defense University, June 17-20, 1997.
[Rat b] A. Rathmell, "CyberWar: The Coming Threat?," *The Newsletter for Criminal Analysts within NCIS*, National Criminal Intelligence Service, July 1997, 2, p.1.
[Ric] M. Richtel, "Study Finds Rise in Computer Crime," *The New York Times*, March 5, 1998.
[Sag] S. Sagan, *The Limits of Safety--Organizations, Accidents, and Nuclear Weapons*, Princeton University Press, 1993.
[Schl] A. Schlesinger, Jr., "Has Democracy a Future?," *Foreign Affairs*, September/October 1997.
[Schw] W. Schwartau, *Information Warfare, Second Edition*, Thunder's Mouth Press, New York, 1996.
[Sta1969] C. Starr, "Social Benefit Versus Technological Risk," *Science*, 168, 1232-1238, September 19, 1969.
[Sta1991] C. Starr, "Twenty Year Retrospective on Risk Management," *Risk Management: Expanding Horizons in Nuclear Power and Other Industries*, 67-72, Hemisphere Publishing Corporation, 1991.
[Ste] S. Steinberg, "Netheads vs. Bellheads", *Wired Magazine*, October 1996, pp.144-147,206-213.
[Tal] Edward B. Talbot, Sandia National Laboratories, Livermore, California, personal communication.
[Ten] D. Tennenhouse, et al, "A Survey of Active Network Research", *IEEE Communications Magazine*, January 1997, pp. 80-86.

[USN] U.S. Nuclear Regulatory Commission, *Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants*, NUREG-1150, Vol. 1-2, U.S. Nuclear Regulatory Commission, December 1990.
[Vau] D. Vaughan, *The Challenger Launch Decision*, University of Chicago Press, 1996.
[Wal] S. Walters, "Learning from Accidental Experiments," *Mechanical Engineering*, p. 44-49, February 1987.
[War] E. Warner, "Eavesdropping is Easy, Modifying Scanners Legal", *Wireless Week*, January 27, 1997, pp. 12-13.
[Wir] P. Wirth, "The Role of Teletraffic Modeling in the New Communications Paradigms", *IEEE Communications Magazine*, August 1997, pp. 86-92.
[You] E. Yourdon and J. Yourdon, *Time Bomb 2000*, Prentice Hall PTR, 1998.
[Zeb1991] E. L. Zebroski, "Lessons Learned from Man-Made Catastrophes," *Risk Management: Expanding Horizons in Nuclear Power and Other Industries*, p. 51-65, Hemisphere Publishing Corporation, 1991.
[Zeb1989] E. L. Zebroski, "Sources of Common Cause Failures in Decision Making Involved in Mad-Made Catastrophes," *Advances In Risk Analysis*, 7, 1989.

M98004739

Report Number (14) SAND--98-0930C
CONF-980433--

Publ. Date (11) 199804
Sponsor Code (18) DOE/CR, XF
UC Category (19) UC-900, DOE/ER

DOE