Cybersecurity and Critical Infrastructure Protection

James A. Lewis
Center for Strategic and International Studies, January 2006

Cybersecurity entails the safeguarding of computer networks and the information they contain from penetration and from malicious damage or disruption. Since the use of computer networks has become a major element in governmental and business activities, tampering with these networks can have serious consequences for agencies, firms and individuals. The question is to what degree these individual-level consequences translate into risk for critical infrastructure.

Analyses of asymmetric, unconventional attacks at first assumed that potential opponents would be drawn to the use of cyber weapons. These opponents could include conventional nation-state opponents and "non-state actors." Cyber weapons were considered attractive for asymmetric attacks because they could offer low-cost means of exploiting the potentially damaging vulnerabilities that are found in most computer networks. Some analysts go further and argue that a cyber weapon could create destruction equal to a kinetic or blast weapon, or could amplify the effects of an attack with these kinds of weapons.

The term "Digital Pearl Harbor" appeared in the mid 1990s, coinciding with the commercialization of the internet. Digital Pearl Harbor scenarios predicted a world where hackers would plunge cities into blackness, open floodgates, poison water supplies, and cause airplanes to crash into each other. But no cyber attack—and there have been tens of thousands of cyber attacks in the last ten years—has produced these results. The dire predictions arose from a lack of insight into the operations of complex systems, from an overestimation of both the interconnectedness of critical infrastructures and the power and utility of software as a weapon to be used against them.

Determining the actual degree of risk posed by computer network vulnerabilities requires an estimate of the probability that a computer malfunction will damage a critical infrastructure in ways that will affect the national interest. For this to occur, a number of simultaneous or sequential events must take place to let a digital attack in cyberspace have a physical effect. This is not a simple transformation. Computer networks are indeed vulnerable, but this does not mean that the critical infrastructures these networks support are equally vulnerable. Terrorists are attracted to different kinds of weapons, particularly explosives, which are more reliable and which better meet their political and psychological need for violence. Infrastructures are robust and resilient, capable of absorbing damage without interrupting operations and accustomed to doing so after natural disasters, floods, or other extreme weather conditions. In short, the cyber threat to critical infrastructure has been overstated, particularly in the context of terrorism.[1]

This initial overstatement does not mean, however, that we should ignore cybersecurity in planning for critical infrastructure protection. First, as the use of computer networks grows, vulnerabilities will increase. Second, a more sophisticated opponent will not use network attacks in an attempt to cause physical damage or terror, but instead target the information stored within computer networks. Nation-states are likely to be attracted to this approach: penetrate networks, collect information and observe activities without arousing suspicion and, should a conflict begin, use that access to disrupt databases and networks that support key activities. This is a different kind of threat from what much of the planning and organization for critical infrastructure protection at first had in mind, and addressing it may require a reorientation of our thinking and our actions on cybersecurity. This chapter discusses reasons and goals for reorientation.

Political Context for Cybersecurity and Critical Infrastructure Protection

There is now a general recognition that cybersecurity was overemphasized in the initial Federal efforts at critical infrastructure protection. Cybersecurity was, at the end of the 1990s, the dominant theme in policy documents and public discussions of critical infrastructure protection.[2]

The overemphasis was the result of several factors. Critical infrastructure came of age in the era when the Internet seemed to have upended all rules. The mentality of the dotcom era underlay many of the assumptions on the scope and linkages of critical infrastructure and cybersecurity. The newness of critical infrastructure protection as an area for security analysis—the U.S. had not contemplated attacks on infrastructure (other than by strategic nuclear weapons) for decades—introduced a degree of imprecision into early analyses. The heightened concern over Y2K, when IT experts warned that ancient programming errors associated with the millennial change would make computers around the world go haywire at the stroke of midnight on New Years and plunge the globe into chaos, helped focus attention on cyber networks as a new and dangerous vulnerability.

Analyses of critical infrastructure protection were also shaped (and continue to be shaped) by a change in American political culture. Evidence for this change is (yet) diffuse and anecdotal, but American government has become progressively more risk-averse since the 1970s. The reasons for this include a loss of confidence among governing elites, decreased public trust of government (with concomitant increases in accountability and oversight requirements) and a more partisan and punitive political environment. The consequences of a more risk-averse political culture are far reaching and have yet to fully play out for the United States, but an exaggerated aversion to risk affects the discussion of strategies for critical infrastructure protection (even if the actual implementation of those strategies is at times lax enough to appear to welcome risk with open arms).

This set of political changes is important for understanding critical infrastructure protection and cybersecurity's place in it. Planning for critical infrastructure protection involves an assessment of risk (the probability that a damaging attack can be made). A risk-averse individual will estimate the probability of a damaging attack as higher than a more neutral approach might suggest. This overestimation of risk has been a standard element of discussions of cybersecurity.

Assessing Risk

Determining the importance of cybersecurity for critical infrastructure protection must begin with an estimate of risk. This has proven to be difficult to do, for some of the reasons suggested above. A neutral approach to estimating risk would look at the record of previous attacks to gain an understanding of their causes and consequences. It would estimate the likelihood of a potential attacker selecting a target and which weapon or kind of weapon an attacker would be likely to use against it (and this involves an understanding of the attackers' motives, preferences, strategic rationale, goals, capabilities, and experience). It would match attacker goals and capabilities against potential infrastructure vulnerabilities, in effect duplicating the analysis and planning process of potential attackers, as they identify targets and estimate the likelihood of success in achieving their goals an attack using a particular weapon and tactics.

The importance of cybersecurity revolves around how we define risk and how much risk a government or society is willing to accept. Homeland Security Policy Directive 7 (HSPD 7), which lays out federal priorities for critical infrastructure protection, begins by noting that it is impossible for the U.S. to eliminate all risk and calls on the Secretary of Homeland Security to give priority to efforts that would reduce risk in "critical infrastructure and key resources that could be exploited to cause catastrophic health effects or mass casualties comparable to those from the use of a weapon of mass destruction."[3] For the purposes of this article, the definition of risk used to assess the need for cybersecurity will be the probability of an outcome that (a) causes death and injuries, (b) affects the economic performance of the United States and (c) reduces U.S. military capabilities.

Using these criteria, there have been no successful cyber attacks against critical infrastructure (much less attacks that produced terror among the population). Even if we use a minimal definition of risk, that an attack results in a disruption in the provision of critical services that harms the national economy and rises above the level of annoyance, there still have been no successful cyber attacks on critical infrastructure.[4]

An even more rigorous approach would limit risk to outcomes that affect the macroeconomic performance of the United States and reduce U.S. military capabilities. Every society has the ability to absorb a certain amount of death and destruction without serious consequence. 2005 saw Hurricane Katrina lay waste to much of the Gulf cost and cost perhaps 2,000 lives (initial and hysterical claims by local officials that Katrina would close New Orleans for years and cost 10,000 deaths were very wrong). Despite the damage and suffering, there was only a small blip in GDP (economists suggest that U.S. economic growth would have reached 4% instead of 3.8% if not for Katrina), and there was no degradation of military capabilities.

One way to estimate the effect a cyber attack is to ask whether a foreign power, using cyber weapons, could stop U.S. military forces from deploying. How, for example, could China prevent a carrier battle group in San Diego or Hawaii from heading for the Taiwan Straits using cyber weapons? Interfering with the telecommunications systems might slow the recall of crew members on leave (if China was able to successfully disrupt the multiple cellular networks in addition to the fixed telecom network and email). Interfering with the traffic signals could make it more difficult for the crews to assemble, as could interfering with the electrical grid, which could also complicate and slow preparing the ships for departure. Hackers could take over broadcast radio and TV stations, to play Chinese propaganda or to change broadcast parameters in the hopes of creating radio interference.

Yet this is a poor start to securing naval victory. If China or another opponent were able to turn off telecommunications, electricity and the traffic light system, it would have little effect on the ability of the carriers to deploy. Further, this sort of attack creates the risk for nation-states (as opposed to non-state actors) of exacerbating tensions or widening conflict in exchange for very little benefit.

The counterargument to these neutral approaches is that they ignore the political effects of a successful attack. The most important of these political effects are the damage to a government's credibility and influence, and the risk of an overreaction by security forces that does more damage than the attack itself.[5] Some scenarios even contemplate non-state actors launching a cyber attack with the knowledge that while its actual effect would be feeble, the overreaction by security forces would be damaging (the history of the Transportation Security Agency and the air passenger business, where large costs to consumers and tax payers are traded for a modest reduction in risk, demonstrates this effect). While the "self-inflicting strategy" may not appeal to violence-prone attackers like Al-Qaeda or other jihadi groups, it is one scenario where the subtle use of cyber attack by a national state could trigger long-term economic damage.

But the political consequences of an attack, cyber or otherwise, can be hard to predict. We know that in many instances, the effect of an attack is to actually harden resistance and increase support for an incumbent government. Even unpopular governments will benefit.[6] Political leaders who put forth the right message of steadfast resolve in the face of attacks will actually improve their standings. While the political investigations that followed September 11 called into question the competence of both the Bush and Clinton Administrations, the immediate political effect was to generate a wave of support for the incumbent President. This support can be lost if the response to an attack is seen as ineffective, but if a government puts forward the right messages, avoids self-inflicted damage, and is seen as making progress in reducing risks of further attacks, any political harm may very well be limited.

Computer Networks and Critical infrastructures

The United States has identified a long list of industries as critical. They include, according to the National Infrastructure Protection Plan, food and water systems, agriculture, health systems and emergency services, information technology and telecommunications, banking and finance, energy (electrical, nuclear, gas and oil, dams), transportation (air, road, port waterways), the chemical and defense industries, postal and shipping entities and national monuments and icons.[7] The nature and operations of most of these infrastructures suggests that cybersecurity is not a serious problem for them.

An infrastructure is judged to be critical because it meets some standard of importance for the national interest—in that the goods or services it provides are essential to national security, economic vitality and way of life. To meet this standard, there is an implicit assumption that the disruption of the infrastructures would reduce the flow of essential goods or services and create hardship or impede important government or economic operations. In the interest of deciding where cybersecurity makes a useful contribution to critical infrastructure protection, we can refine this standard by introducing two additional concepts—time and location.

Time and location help explain why cybersecurity is not of primary concern for many critical infrastructures. If there are immediate problems when a system goes off-line, not problems that emerge after weeks or months, that system is critical. Problems that take longer to appear allow organizations to identify solutions and organize and marshal resources to respond, and thus do not present a crisis. The ability of industrial societies to respond to problems, to innovate and to develop alternative solutions or technologies, suggests that in those infrastructures where disruption does not produce immediate danger and was not prolonged for an unreasonable period of time, there would be little effect on national security, economic vitality or way of life.

There is also a geographic element to criticality. National infrastructures are composed of many local pieces, not all of which are equally critical. Specific elements of the larger infrastructure provide critical support to key economic and governmental functions, not entire networks or industries. It is harsh to say, but Hurricane Katrina in 2005 demonstrated that large cities or sections of the country can be taken offline and, if the political consequences are managed, have little effect on national power—either economic or military. Certain high-value targets—the national capital region, military facilities, a few major cities, or nuclear power plants—require greater attention across the board, while other places, where disruption or destruction would not impair key national capabilities, can be assigned a lower priority.

The concerns of cybersecurity can transcend this geographic focus in some instances. There are a few, very few networks that are national in scope and interconnect thousands of entities in ways that make them mutually dependent. However, these networks—finance, telecommunications, electrical power—are among the most critical for national security and economic health, and their interconnectedness, national scope and criticality may make them more attractive targets for cyber attack.[8] Fedwire, the financial settlement system operated by the Federal Reserve Banks, provides a crucial service to banks. Interfering with Fedwire would cripple (temporarily) the U.S. banking system. The Federal Reserve has expended considerable effort to harden FedWire, and the Fed's desire to prevent online bank robbery provides an incentive to continue these efforts.

The U.S. electrical system is composed of several thousand public and private utilities organized into ten large regional grids. There is a substantial degree of interconnection within these grids and computer networks play an important role in managing grid operation and the production of electrical power. The grids themselves suffer form the consequences of underinvestment and deregulation. Newer industrial control systems use commercial computer operating systems and IP protocols as they are cheaper and easier to use. However, the new technologies replace older control systems that used with specialized proprietary software and dedicated networks that were difficult for hackers to access and exploit. The move to commercial software and IP increases vulnerability.

Vulnerability is not the same as risk, however, and a number of factors limit the increase in risk created by this transition to "off-the-shelf" control systems. There have been thousands of hacking incidents aimed at power companies, but as of yet, none have produced a blackout.[9] In the larger national context, blackouts are common in the U.S. and often do not even attract national attention. In 2002, an ice storm blacked out the 20th largest city in the U.S. with a population of 600,000 for several days. The event had no effect on economic or military power and barely merited attention in the national press. Power companies cooperate to respond quickly to these events. Many critical facilities have installed backup power generation equipment. A localized blackout outside of few major cities can be of minor importance to the nation—witness the recent Los Angeles blackout. The real risk may lie in interconnection, and the ability of an attacker to access one vulnerable producer and cascade this attack into a blackout of one of the big regional grids, but an attack that succeeds in blacking out a single facility might only be seen as an annoyance.

Telecommunications services are another national-level network. The telecom backbone that supports the internet and voice communications is comprised of a number of large networks. An attack that disrupted the services provided by several of these large networks could disrupt communications traffic. However, the presence of multiple overlapping connections means that there is no single point of failure. The use of satellites in communications services also introduces a degree of redundancy. Since the 1970s, telecommunications networks have been hardened to allow for some continuity of service even after a strategic nuclear exchange.

Additionally, telephone companies developed and use packet switching technology (which breaks messages into many small "packets" of data that can be sent separately) to allow voice communications to persist without a continuous end-to-end connection. The internet relies on packet switching and benefits from the robustness provided by this technology. The internet itself was designed to automatically route around damage to complete transmissions. Communications may be slowed or disrupted, but there is no single point to attack that would easily allow the national telecommunication system to be disabled.

Before deregulation and the breakup of the national monopoly, the U.S. telecom network was built (with Federal guidance) to provide survivability and redundancy in the event of attack, accident or system failure. After deregulation, when telecom companies were less able to make investments solely to meet the requirements of national security, a highly competitive environment and rapid technological development became the source of a high level of redundancy. In contrast to an attack that destroyed facilities, a cyber attack would (a) require sustained, successful re-attack to overcome network operators' repair efforts and (b) would have to disable multiple communications systems (wireless, fixed line, internet) to degrade communications.

The complexity of successfully carrying out a cyber attack against national infrastructures like telecommunications or the electrical grid, combined with a lower probability of success than a physical attack, may make it unattractive to terrorists. Terrorists want screaming people to run in terror past mangled bodies in the street—an attack that only produces a busy signal is likely to be dissatisfying. In theory, the idea of a cyber attack against telecommunications systems in coordination with a physical attack is attractive, as it could compound damage and terror, but coordinating two simultaneous attacks adds a degree of complexity that may overwhelm a terrorist cell's planning capabilities while increasing the chances of detection.

The same constraints do not apply to a nation-state attacker. Such an attacker would have the resources for coordinated attacks. Surreptitious economic warfare during peacetime may be attractive, but an opponent would want to weigh the benefits of an attack that produced a long-term drag on the target's economy against the risk and damage of discovery. In the event of a conflict, however, a nation-state opponent is likely to use cyber weapons to attempt to disrupt these large U.S. national networks.

The Internet as a Critical Infrastructure

Some point to the Internet as a single large infrastructure that could be attacked with cyber weapons.[10] The first point to bear in mind, however, is that it is a shared global network. An attack against it will affect both target and attacker. An attacker may calculate that the U.S. might suffer more as a result, or it could plan to use some alternative or backup system to replace the internet while the target struggled to respond, giving it a temporary advantage.

The internet is very robust. It is a network designed to continue to function after a strategic nuclear exchange between the U.S. and the Soviet Union. Its design and architecture emphasize survivability. The internet (building on earlier technological improvements created by packet switching in telecommunications) could deal with disruption by automatically rerouting to ensure that a message would arrive despite the complete destruction of key nodes from the network. The internet addressing system, which is critical to the operations of the system, is multilayered, decentralized, and can continue to operate (albeit with slow degradation of service) even if updating the routing tables that provide the addressing function is interrupted for several days. Some of the core protocols upon which the internet depends appear vulnerable to attack. BGP (Border Gateway Protocol) is responsible for routing traffic and a number of tests suggest that BGP is vulnerable to attack but an attacker faces the immense redundancy contained in a network comprised of tens of thousands of subsidiary networks.

There has been at least one effort to attack the Internet. An October 2002 attack by unknown parties used a Distributed Denial of Service attack against the 13 “root servers” that govern Internet addresses. The attacks forced eight of the thirteen servers off-line. The attack on the DNS system did not noticeably degrade Internet performance and went unnoticed by most of the public, but had it been continued for a longer period (and if the perpetrators remained undetected) there could have been a significant slowdown in traffic. A successful attack on the Internet’s DNS system, if successful, would slowly degrade that system’s ability to route traffic, but this would take several days to have any effect. In response to the attack, the DNS system has been strengthened since the 2002 attack by dispersing the root servers to different locations in and by using new software and routing techniques. The new redundancy makes shutting down the DNS system a difficult task for an attacker.

The difficulty of estimating the actual cost of a cyber attack adds complexity to planning for critical infrastructure protection. Estimates of damage from cyber attacks at times reflect the heritage of the dot.com boom in cybersecurity—they generally overestimate or exaggerate damage. Damages are estimated by taking a sample of costs to various users and then extrapolating them to the affected user population. In some cases, the sample of costs is itself an estimate. These estimates of the economic damages of cyber attack show considerable variation in the value they ascribe to cyber incidents. There is also considerable variation in their methodologies, which are often not made public. Few if any of these efforts use the sampling techniques derived from statistical analysis that could ensure greater reliability. Statements that cybersecurity is crucial because of the risk of economic losses that could total in the millions, hundred of millions or billions of dollars should not be accepted at face value.

It is important to disaggregate the effects of an attack. Analysts often cite the Slammer worm as a damaging cyber attack, but its effects were, from a national perspective, inconsequential. One frequently cited example about the damage of Slammer tells how it affected automatic teller machines (ATM) across the northwest, putting 13,000 of them out of service. What is important to note, however, is that Slammer affected only one bank and its ATM network. Other banks were unaffected, and the other major bank in the region did not see its AMT network go offline at all. In this instance, customers of the first bank were inconvenienced. The first bank lost revenue and suffered reputational damage. The bank’s competitors were, in one sense, rewarded for practicing better cybersecurity, as some transactions that would have been made on the first bank’s ATM network were instead conducted on their machines.

Another example involves a railroad forced by the ‘sobig’ virus to suspend operations on 23,000 miles of track—but no other railroad was forced to suspend operations.[11] If a cyber attack damages one company in a critical sector but leaves its competitors operational, it limits the overall risk to critical national functions. It is difficult to think of a case where a cyber attack affecting one firm and not others would pose a risk to security.

We do not want to extrapolate the misfortunes of a single company to an entire sector in estimating the risk to critical infrastructure from a cyber attack. Similarly, we also want to disaggregate the estimates of opportunity cost to determine whether it is a single company that suffers or the entire economy. In this case, opportunity cost refers to the income (or production) lost when a resource cannot be used, a sale made, or a service provided because of cyber attack. Most of the estimates of the cost of the damage of cyber attacks include an estimate of opportunity cost and this often makes up a large portion of the estimated damages from an attack.

Opportunity cost can be misleading for security analysis. If one online merchant is forced offline by a cyber attack, but their competitors remain in operation, customers may choose to go to the site that works to make their purchase. The vendor forced offline has lost income, but national income remains unaffected. Other customers may choose to wait and make their purchase later. Again, national accounts are ultimately unaffected. In other cases, a manufacturer may see its website or corporate email network go offline but be able to maintain production—in one case where a virus damaged an auto manufacturer's corporate email systems, the production for cars and trucks was unaffected.[12]

This is not to disparage the effects of cybercrime, which can be costly for an individual or company. However, most cybercrime involves losses in the thousands of dollars (there are anecdotal reports that a few major banks have experienced much larger losses, but they have not made these losses public for fear of reputation damage). Cybercrime is prevalent and increasing, but this does not mean that the risk to critical infrastructure is similarly increasing. Cybercriminals want money. Their favored tactics include theft of valuable data or extortion (e.g., the threat to launch denial of service attacks or disrupt networks unless paid). Their first question will be how to turn a threat to a national infrastructure into financial gain without risk of arrest. Threatening an attack against multiple firms may either be operationally too difficult or attract too much attention from law enforcement agencies.

In view of their motives and incentives, attacks against infrastructure by cybercriminals seem unlikely, but nation-states may adopt the hacking tools and “bot nets” developed by cybercriminals for use in cyber war efforts. The sophisticated “shareware” available for cybercrime from hacker or “warez” sites can give even inexperienced attackers access to advanced automated tools and techniques. These range from online hacking manuals and do-it-yourself virus kits to sophisticated attack tools that require some computer expertise to use.

The most interesting of these tools allow a hacker to place surreptitiously malevolent programs on a computer without the user’s knowledge. The program can then execute damaging instructions, transmit data or an external address, or provide increased (and invisible) access and control to the hacker. This “malware” can infect computers through the opening of malicious e-mail attachments, downloading seemingly harmless programs, or simply through visiting a malicious Website. Cybercriminals assemble networks of these infected computers for use in denial of service attacks, for spamming or for advertising and tracking. Using these tools, an attacker could attempt to disrupt networks and damage or erase data.

However, not all networks are equally vulnerable to the tools of cybercrime. The botnets mainly infect consumer systems using always-on connections. Damaging these consumer computers would be annoying, but not threatening to national security or long-term economic health. Second, cybercrime tools are not aimed at physical infrastructure. Infecting a computer does not automatically become a risk to an associated infrastructure. This means that while cybercrime can increase, and it is a growing problem for law enforcement, that does not mean that the risk or critical infrastructure is also increasing.

One benefit that has come from the attention paid to cybercrime is that the measures that improve cybersecurity to protect against criminals will also reduce any risk to critical infrastructure. The use of regular software patching, defensive software (such as intrusion detection systems and anti-virus and anti-spyware software), better authentication of users and encryption of sensitive data will make an attackers job more difficult. Improved law enforcement capabilities to arrest and prosecute cyber criminals will also reduce the attractiveness of cyber attacks against critical infrastructure. Companies are more likely to spend money to protect themselves from criminal attacks, since this offers a direct and immediate benefit their bottom line. Defense is a "public good"[13] and the private sector routinely undersupplies public goods. This is a particularly important point given the U.S. dependence on the private sector for critical-infrastructure protection. A reading of economic incentives suggests that companies will spend to improve cybersecurity to prevent cybercrime more than they would for a nebulous threat to homeland security.

The importance of cybersecurity in protecting critical infrastructures other than finance, electrical power or telecommunications, rests on the assumption that critical infrastructures are dependent on computer networks for their operations. The chief flaw in this reasoning is that while computer networks are vulnerable to attack, the critical infrastructures they support are not equally vulnerable. Early proponents of cyber attack assumed that many public services, economic activities and security functions were much more dependent on computer networks than they are in their actual operation. While the dependence on computer networks continues to grow, many critical functions remain insulated from cyber attack or capable of continuing to operate even when computer networks are degraded. It may be more accurate to say that critical infrastructures are dependent on their human operators, whose actions are supported, reinforced or carried out using computers and networks. This human element reduces the risk of cyberattack to critical infrastructures.

A well-known example of the difference between computer vulnerability and system vulnerability again comes from the 2003 release of the Slammer worm. The worm affected database software. Some police departments in Washington State saw the computers used in their 911 emergency response systems slow to the point of uselessness as the worm spread and implemented its instructions. These departments compensated by using paper notes to record calls, allowing 911 services to continue uninterrupted.[14] The computers were vulnerable and affected, but the critical service was not.

The debate today in how to approach this task is whether cybersecurity should be an element of a larger critical infrastructure strategy or whether it deserves its own independent approach. While the first phase of planning for critical infrastructure protection made cybersecurity of primary importance, the second phase of thinking about critical infrastructure protection assumed that cybersecurity only made sense as part of a larger strategy focused on physical protection. Had the 911 centers in Washington State been the subject of physical attack or damage, they might very well have had to shut down and 911 services would have been disrupted (as was the case in New Orleans the post-Katrina flooding). Incidents like this seem to show that risk comes primarily from physical attack.

While this new approach to critical infrastructure protection has dominated federal planning for the past few years, it is not universally accepted. There are reasons for this lack of universal acceptance, some good, some less sensible. The IT industry did not like being downgraded from the central place it occupied in critical infrastructure protection. A more cogent argument for a separate approach to cybersecurity involves recognition of the inability of differing security communities to implement a strategy that unifies cyber and physical security. A Chief Security Officer in a corporation often thinks in terms of “guns, gates and guards.” The Chief Information Security Officer thinks in terms of firewalls and software. In most companies, neither is well placed to execute a unified strategy.

A third approach to critical infrastructure protection might be to recognize that the importance of cybersecurity varies from infrastructure to infrastructure and with the nature of the attacker. We should not be surprised that the distribution of vulnerability is not uniform among or across infrastructures. Cybersecurity is more important for a few networked, interconnected national infrastructures and less important for many disaggregated infrastructures. Cybersecurity is less important in planning to defend against terrorist attacks, since these are less likely to use cyber weapons, but more important in planning for conflict with a national state opponent. Critical infrastructure protection could distinguish among the many places where cybersecurity is a tertiary source of risk and the few places where it is of central importance. These key facilities could be “hardened” with a combination of redundancy, contingency plans for responding to computer disruption, maintaining non-networked controls for key functions, and by ensuring additional monitoring of computer and network activities.

Conclusion

HSPD-7 asserts “Terrorists seek to destroy, incapacitate, or exploit critical infrastructure and key resources across the United States.” This assertion is not entirely accurate. While terrorist do exploit western infrastructure for transport and communications to obtain a global presence and capability, it is not clear that they seek to destroy or incapacitate critical infrastructures. Their strategies do not emphasize economic warfare, but favor a blend of military and psychological actions that they believe will produce political change. Cyberspace is a valuable tool for coordination and propaganda for terrorists, but it is not a weapon.[15]

Nation-states who are potential opponents may see more opportunity in cyberspace. Intelligence gathering will prompt them to penetrate U.S. computer networks. In the event of a conflict, nation-states will likely try to use the skills and access gained in intelligence operations to disrupt crucial information systems. This disruption will also affect critical infrastructures and, potentially, degrade the services they provide. It remains unclear, however, if even a skilled opponent can translate the degradation of key infrastructure services into military advantage for a conflict whose combat phase is likely to be of short duration and depend more on existing inventories.

The best path to better cybersecurity may lay outside of critical infrastructure protection. It is hard to motivate people to defend when risks are obscure or appear exaggerated. However, the risks of espionage (including economic espionage) and cybercrime are very real for individuals, firms and agencies. A security agenda that focused on measures to respond to cybercrime and espionage would produce tangible benefits, win greater support, and reduce much vulnerability in computer networks used by critical infrastructure. If an emphasis on cybercrime and counterespionage is the key to better cybersecurity, this suggests that the core of the problem lies with law enforcement.

Critical infrastructure protection began by making cybersecurity the cornerstone of defense. This chapter suggests that in fact, if we calculate the risk from cyber attack for most infrastructures, it is a tertiary concern. The history of critical infrastructure protection has been to develop expansive plans to cover a broad list of targets and then, in the effort to protect many things with few resources, achieve little in terms of risk mitigation. Putting cybersecurity in the context of more precise assessments of the actual threat could help overcome some of this difficulty by allowing a federal strategy to focus on the few networks of real concern.

About the Author

James A. Lewis is a Senior Fellow and Director for Technology and Public Policy at the Center for Strategic and International Studies (CSIS), a research institution in Washington. Before joining CSIS, Lewis was a career diplomat at the Department of State, and a member of the Senior Executive Service at the Department of Commerce. He received his Ph.D. from the University of Chicago in 1984.

Notes

[1] “Assessing the Risks of Cyber terror, Cyber war and Other Threats” provides the fuller discussion of the likelihood of cyber terrorism. http://www.csis.org/media/csis/pubs/021101_risks_of_cyberterror.pdf

[2] The Joint Security Commission was the first in a series of commissions to identify cybersecurity as a primary challenge, saying "The Commission considers the security of information systems and networks to be the major security challenge of this decade and possibly the next century…” Joint Security Commission, “Redefining Security: A Report to the Secretary of Defense and the Director of Central Intelligence,” February 28, 1994, Chapter 1, “Approaching the Next Century”

[3] “Homeland Security Presidential Directive/HSPD-7: Critical Infrastructure Identification, Prioritization, and Protection,” December 17, 2003, http://www.whitehouse.gov/news/releases/2003/12/20031217-5.html

[4] A fuller discussion of this claim, and use of the concept of ‘opportunity cost’ in assessing economic harm, follows below.

[5] This conclusion reflects the results of government-sponsored cyber war games.

[6] The first study to confirm this counterintuitive effect was the U.S. Strategic Bombing Survey, but later studies have found a similar reaction among target populations. U.S. Strategic Bombing Survey, Summary Report (European War), 1945. See also Stephen T. Hosmer, “Psychological Effects of U.S. Air Operations in Four Wars,” Rand, 1996

[7] The earlier PDD-63 (May 1998) identified the task as protecting “the nation's critical infrastructures from intentional acts that would significantly diminish the abilities of: the Federal Government to perform essential national security missions; and to ensure the general public health and safety; state and local governments to maintain order and to deliver minimum essential public services; and the private sector to ensure the orderly functioning of the economy and the delivery of essential telecommunications, energy, financial and transportation services.” The Patriot Act and HSPD 7 also provide similar but not identical lists of infrastructures deemed critical.

[8] Oil and gas pipelines could be considered a national network, but there are alternative transport modes that could mitigate an attack. Air traffic control may appear national, but is conducted in discrete segments on a local and regional basis.

[9] "Energy and power companies experienced an average of 1,280 significant attacks each in the last six months, according to security firm Riptech Inc…. The number of cyber attacks on energy companies increased 77 percent this year (2002)." CBS News, "Hackers Hit Power Companies, July 8, 2002, http://www.cbsnews.com/stories/2002/07/08/tech/main514426.shtml

[10] For more on this, please see the chapter by Aaron Mannes in this volume.

[11] National Infrastructure Advisory Council, "Prioritizing Cyber Vulnerabilities," October 2004, Page 5, at http://www.dhs.gov/interweb/assetlibrary/NIAC_CyberVulnerabilitiesPaper_Feb05.pdf

[12] The Ford Motor Company received 140,000 contaminated e-mail messages in three hours. It was forced to shut down its email network. E-mail service within the company was disrupted for almost a week. Ford reported, "the rogue program appears to have caused only limited permanent damage. None of its 114 factories stopped production. Computerized engineering blueprints and other technical data were unaffected. Ford was still able to post information for dealers and auto parts suppliers on Web sites that it uses for that purpose." Keith Bradsher, "With Its E-Mail Infected, Ford Scrambled and Caught Up," The New York Times, May 8, 2000

[13] A "public good" provides benefits to an entire society with very little incentive for any one person to pay for it.

[14] Wells, R. M., "Dispatchers go low-tech as bug bites computers" Seattle Times, January 27, 2003, http://archives.seattletimes.nwsource.com/cgi-bin/texis.cgi/web/vortex/display?slug=webworm27m&date=20030127

[15] See, for example, Office of the Director of National Intelligence, "Letter from al-Zawahiri to al-Zarqawi, October 11, 2005," http://www.dni.gov/letter_in_english.pdf