Cyber Threats to the Financial Sector:

Understanding the Attack Surface

Thomas Garwin, Nicholas Rostow and Abraham Wagner

December 11, 2023

MARGIN RESEARCH

The research described in this report was sponsored by the Defense Advanced Research Projects Agency (DARPA). The views expressed are those of the authors and do not necessarily reflect the views of the U.S Government.

www.margin.re

Contents

Foreword and Acknowledgements

What began as a research project at the Advanced Research Projects Agency (ARPA), as it was then known, in the 1960s evolved into a technology revolution never anticipated as these new technologies gave rise to the most significant paradigm change since the invention of movable type in the 15th century. Now most communications and information operations now take place on systems connected to the Internet with all major sectors including finance becoming highly dependent on this critically important infrastructure.

Dependencies on the Internet infrastructure make users vulnerable to both criminals and potential state adversaries. While ARPA's early efforts involved only networking scientists and did not take security into account, the ARPAnet and then the successor Internet were insecure and prone to hostile attacks of all kinds. Together with espionage and cyberwarfare, cyber attacks have become a substantial national security problem and a major threat to the financial sector.

The most significant foreign cyber threats now come from China, Russia, Iran and North Korea, as these states recruits skilled personnel; develop malicious code; and prepare for hostile cyber operations. They have greatly expanded their cyber capabilities in intelligence collection, espionage, deception, and cyber warfare where the prospect of a major attack or "Digital Pearl Harbor" could devastate the nation's financial sector in a non-kinetic attack.

The present analysis considers the vulnerability of the financial sector and the attack surface involved. It also considers various options for increasing the resilience of the financial sector to such a cyber attack and responding to it.

This study effort was made possible with support from the Defense Advanced Research Projects Agency (DARPA). The study team has benefited greatly from discussions with personnel from the Treasury Department, other U.S. Government agencies, as well as various individuals within the financial sector. Also supporting the work have been several research assistants, currently graduate and law students at Harvard University and New York University (NYU) School of Law. The views expressed do not reflect the views of any organization or the U.S. Government.

Executive Summary

The initial ARPAnet of the 1960s evolved into a technology revolution never anticipated, giving rise to the most significant paradigm change since the invention of movable type in the 15th century. Most communications and information operations now take place on the Internet infrastructure that supports the financial sector and others. Critical sectors have become increasingly dependent on this important resource, making users vulnerable to both criminals and potential state adversaries.

While early efforts involved only networking scientists and did not take security into account, the original ARPAnet and the successor Internet were insecure and prone to hostile attacks of all kinds. Cybercrime, espionage, and cyberwarfare have become a major national security problem with the most significant foreign cyber threats coming from China, Russia, Iran and North Korea. These states recruit skilled personnel, develop malicious code, and prepare for hostile cyber operations, greatly expanding their cyber capabilities. The prospect of a major attack or "Digital Pearl Harbor" could devastate the nation's financial sector in a non-kinetic attack.

Government, industry, and academia have all recognized the growing threats but have largely failed to address the most serious aspects of the problem. The financial sector has largely focused on "low hanging fruit" and far less on the capabilities being developed by potential adversaries for a major cyber attack on specific systems and technologies. For decades, the U.S. has been complacent about this risk of major cyber attack, and the idea that the commercial sector and private industry would solve many cyber problems has proved invalid. The much touted "public-private partnership" likely will not be totally effective.

At the same time responsible government agencies have tended to focus on "incident reporting" rather than detecting hostile code before it is used and preventing a major attack in advance. Such an approach leaves open the prospect of a major intelligence failure with catastrophic consequences.

The present context requires that the nation monitor the development of hostile cyber capabilities and the creation of malicious code as well as the institutions and individuals involved. Reporting cyber attacks as "incidents" after they take place is no substitute for understanding how a devastating cyber attack might be conducted and managed if not prevented.

The Digital Revolution and the Internet

As a result of the digital revolution, digital data has largely replaced analog files and other media. Most of the world's communications and information technology systems have become part of the "connected world." Over the last 30 years the Internet and reliant information technologies have become increasingly central to all aspects of life. At the dawn of the Internet age there was a sophisticated view that large businesses – especially the large money center banks – would in their own interest somehow force Internet technology suppliers to provide inherently

secure products. While many of these products are useful in dealing with lower level attacks, they are not effective against the types of major hostile attacks being developed.

The public and to a major extent the U.S. government have both been slow to recognize the increased danger from the cyber capabilities of China and Russia in the context of dramatically increased antagonism. These autocratic countries have again become vociferous critics of the rules-based international geopolitical, monetary, finance, and trading regimes that the U.S. has championed since World War II.

The Expanding Attack Surface

Critical sectors all became reliant on networked digital services that provided a new venue for crime, espionage, and warfare. Unlike kinetic warfare, cyber threats differ in terms of attribution as well as scale – ranging from the annoying and inconvenient to cyberwarfare that could disable much of the nation's operations in catastrophic cyber-attacks. The Internet today is a far different place than it was in the early years. It is essential to assess the threat landscape and the range of actions that can enhance the resilience of the financial sector where virtually all records are electronic, using systems that are all network-connected.

A major attack by a hostile state actor that resulted in the loss of service to key elements of the financial sector would seriously cripple and likely halt the economy. Safeguards built for the "brick and mortar" world have little or no relevance to transactions in the cyberlandscape. Cyber attacks do not require tactical coordination or military risking their lives, and there is little downside to failed attempts at penetrating critical systems. They can scale from lower level espionage or covert attacks to major ones that can disable the financial sector – giving the attacker a far wider set of options.

Deterring Cyber Attack

Deterrence has long been fundamental to U.S. national security strategy. Potential adversaries, including nation states and non-state actors such as terrorist groups and criminal enterprises, have been deterred from attacking the U.S. and allied nations because of the unacceptable costs from retaliation, both conventional and nuclear. As the range of possible attacks now includes cyberwarfare, policy and strategy for deterrence is needed to incorporate this new domain of espionage and warfare. This remains an essential pillar of the nation's overall cybersecurity strategy.

Problems with deterrence mean that protection of the financial sector from attack, and rapid restoration in the event of attack damage, are crucial. China's technical capacity and position in many supply chains means that managing the U.S.-China trade and political relations and trade and technology interdependencies will also likely be key to avoiding disastrous geopolitical escalation.

Technology Trends

The geopolitical character of cyber threats has changed, and not many fully understand how changes in technology and business complicated efforts to secure computer networks and data systems against sophisticated attacks of the type that China and Russia are equipped to execute.

Further, foreseeable technological developments could pose even more problems. Here some of the technology trends already affecting cybersecurity operations include:

- Increased requirements for systems to be accessible and operated remotely.
- Increased overall complexity of software systems, and reliance on elements which can have vulnerabilities introduced by malicious actors.
- "DevOps" rapid software improvement cycles and increased diversity of hardware and software platforms creating vulnerability routes for malicious access.;
- Increased complexity of software supply chains and maintenance by third parties.
- Increasing reliance on software execution "in the cloud" – on servers not owned or controlled by the corporation.

These recent technological trends will soon be joined by other complications, including:

- Use of Large Language Models (LLMs) to write and patch code.
- Use of chatbots and deepfakes for phishing and other social engineering attacks.
- Quantum computing or other breakthroughs.

Vulnerability of the Financial Sector

The financial sector is a critical component of the nation's infrastructure, and includes banks, securities exchanges, as well as providers of the key financial systems and services. Together they comprise some $108 trillion in assets and faces a variety of cybersecurity-related risks.

There is the potential for monetary gains from criminal activities undertaken against the sector. Apart from criminal and monetary gains there is a more serious concern an advantage to hostile actors from economic disruptions and harm to the nation that can be accomplished without a kinetic attack. This increases the financial services sector's attractiveness as a target for malicious actors. Key risks include:

- Increasing access to financial data through information technology service providers and supply chain partners.
- Growth in sophistication of malware, and other malicious software.
- Greater interconnectivity through the use of networks, cloud service providers, and mobile applications.

As the world has moved to largely cashless operations these institutions maintain all records online; their customers use their web sites and applications for most transactions; card systems; as well as automated clearing house (ACH) systems that are all vulnerable to hacking.

The U.S. financial sector has already devoted substantial attention and resources to cyber security, with the emphasis on restoring operations. This concern, however, may not be adequate in the face of Chinese and Russian capabilities in the context of a geopolitical crisis or hot war where this would be a non-kinetic collateral attack.

Even against normal "peacetime" cyber threats the data suggests that little progress has been made in reducing losses or reducing the time to recognize an intrusion or fix the problem after an intrusion is recognized. Ensuring the cyber security of the financial sector requires both greater vigilance in the near term and use novel technologies to increase the resilience of IT systems in the longer term, It also needs a robust effort to counter cyber intrusion and attack threats. Thirty years of history show that financial institutions and the supporting information technology companies will not achieve this on their own.

Reporting cyber attacks as "incidents" after they take place is no substitute for understanding how a devastating cyber attack might be conducted and managed if not prevented. The country needs to take a far more disciplined approach to dealing with the threat, including detection of malicious computer code development, as well as those engaged in these hostile activities, so that the attack surface is made less vulnerable, and critical systems are made more resilient.

The Impact of Crypto-Currency

Increasingly crypto-currency exchanges and block chains can also serve as stores of value and transactions processing alternatives to banks but these still play a small role in the overall economy. Unfortunately, crypto-currency and the already significant decentralization and diversity of the American financial system are not necessarily a strength in terms of resistance and resilience in the face of cyber attack. Smaller banks have less than state-of-the-art defenses and restoration plans. They are also dependent on a few large financial and information processing vendors and processing relationships with a few credit card networks and a few money center banks.

Depository institutions lend multiples of their deposit base and are subject to bank runs if consumers and others concerned about access to their money. As a result, problems in one bank may cause problems for otherwise unaffected banks. The diversified structure of the banking system may in fact produce a situation where interdependencies between less- and more-vulnerable institutions add vulnerability overall rather than making the system more robust.

Key Role of the Federal Reserve System

All U.S. banking and related services rely on the monetary and payment system run by the Federal Reserve System, including the Federal Reserve Board and the 12 Federal Reserve Banks. The Federal Reserve System:

- Conducts the nation's monetary policy.
- Promotes the stability of the financial system, as well as the safety and soundness of individual financial institutions, monitoring the financial system as a whole.
- Fosters payment and settlement system safety and efficiency.
- Promotes consumer protection and community development through supervision and examination, research and analysis of emerging consumer issues and trends, community economic development activities, and administration of consumer laws and regulations.

In addition to the Federal Reserve, other government and international entities have key roles affecting financial sector operations:

- The Treasury Department has a huge presence in financial markets through its operations to fund the Federal government debt.
- The FDIC and the Office of the Controller of the Currency, play key roles in supervising retail banks and thrift institutions.
- Investment activities are regulated by the Securities Exchange Commission and the Commodity Futures Trading Commission.

Current and Emerging Technologies and Techniques to Reduce Cyber Risk

Measures to enhance cyber security and resilience involve the "CIA triad" of data protection – protecting the Confidentiality, Integrity, and Availability of data and data processing systems. Attacks may target and affect one, two, or all three of these dimensions. Compromises of confidentiality can result in irremediable damage as response may be equivalent to closing the barn door when the horse is gone. Damage to data integrity can be repaired if the information required to do so is available. Denial of service attacks are inherently reversible in terms of the information technology system but can have irreversible consequences in the real world.

Broadly speaking, current and emerging efforts can usefully be divided into several categories:

- Static Defense
- Reducing Vulnerabilities in Deployed Systems
- Resilient Operation Under Attack

These categories broadly reflect the evolution of approaches to cyber security. Within each category there has been some progress largely in response to threats becoming more sophisticated. Some technological and topological trends are common across categories, including increasing use of artificial intelligence techniques beyond rule application, including machine learning, and roles for service providers that are positioned to monitor a wide swath of public and private network activity to discern new threats and vulnerabilities.

Threat Model – Why the Nation-State Threat is Different

Strangely, the history of hacking can be traced to the late 1800s when computer systems didn't yet exist, although there are enough similarities to the early days of exploits to compare telephone line switching and modern cybersecurity. In the current world there are numerous examples of hackers using telephones to create exploits in technical systems. By the mid-1960s, as modern systems evolved, hacking as it is now known actually began. The practice later became associated with changing the technical aspects of a computer to alter it from its intended usage.

Agencies started to test their network security in response to early computer hacking, and by the 1980s which prompted Congress to enact the Computer Fraud and Abuse Act (CFAA). As computers became more widespread and mainstream so did hacking. The late 1990s saw a rise in hacking, as credit card fraud and illegal wire transfers boomed. Cybersecurity became increasingly

significant with increased reliance on computer systems, the Internet and wireless network standards such as Bluetooth and Wi-Fi, as well as the rapid growth of smart devices and the various connected devices in the "Internet of things."

Chinese Cyber Operations

Chinese leaders have long valued access to technology and information to support their national objectives and military capabilities. The Chinese Communist Party (CCP) has always understood the importance of controlling information to maximize its ability to manage competition and conflict. Starting in the 1970s, China moved to acquire technologies in order to collect, store, process, and manage information. Now their success in cyber and communications development is most visible in areas such as 5G (communications) and artificial intelligence (AI).

As part of their long-term competition with the U.S., the Chinese view collection and hoarding of information as an investment in the future. It is a strategic aim, not merely a near term tactic. In the area of cyberwarfare China looks at cyberspace in the broader context of information space. The ultimate objective is, not "control" of cyberspace, but control of information, a vision that dominates China's cyber operations. Chinese military strategists have also begun to discuss the emergence of what they refer to as "intelligentized warfare," which includes the use of information analysis and AI technologies to target an adversary's "cognition."

Having invested heavily in technologies related to surveillance, espionage and cyberwarfare China's cyber operations reflect major advances, and have reshaped their national cyber ecosystem. The PRC applies an integrated Military-Civil Fusion (MCF) strategy as a way to greatly expanded China's cyber capabilities in intelligence, espionage, deception, and cyber warfare.

China's cyber capabilities and operations have increased exponentially to the point where they pose a highly significant national security threat to the U.S. and all perceived adversaries. The size of the attack surface exacerbates the risk from such cyber operations and capabilities. China continues to invest heavily in this technology path and the threat will continue to grow. All U.S. institutions need to assess their vulnerabilities and manage their risks in light of what China is doing.

Russian Cyber Operations

Russian use of information technology and the Internet has strong historical roots. For more than a century, Russia has used forgeries, disinformation, and falsehood-propagation as an important military and intelligence tools. In particular "active measures" are a key element, with covert and deniable political influence and subversion operations, as activities include corruption and disinformation to assassination and sponsorship of coups. Since the time of the Tsars has Russia has used "active measures" to spread false information. Now the Russians emphasize deniability, blur the lines between public diplomacy and propaganda, and use disinformation as a form of political warfare.

The Putin regime also believes in using asymmetric tactics, such as executing assassinations and running disinformation campaigns, to achieve its aims at home and abroad. The Russian government now sees the Internet and the free flow of information it engenders as both a serious

threat and equally serious opportunity. Russian military theorists avoid using the terms cyber or cyberwarfare, preferring to see cyber operations in the broader framework of information warfare. This holistic concept includes computer network operations, electronic warfare, psychological operations, and information operations. Russia views the struggle over "information space" as constant and unending.

Moscow also views cyber operations as means of disruption for disruption's sake. It can degrade an enemy's military communications, disrupt a foreign company's operations in Russia, and achieve other objectives by means that do not amount to an overt use of military force and still retain the ability claim plausible deniability.

Offensive cyber operations therefore play a large and increasing role in Russian military operations and strategic deterrence. While the Russian military and intelligence services were slow to embrace cyber operations, the government has made significant investments in the last decade and continues to bolster offensive and defensive cyber capabilities. Russian patriotic hackers, front groups, and cyber-criminal syndicates, added to military and intelligence capabilities, have become central to Russian offensive cyber operations.

Patriotic hackers and criminal networks have been augmented, if not entirely replaced, by the FSB (the Russian Federal Security Service and successor to the Soviet KGB), and the GRU (Russian Military Intelligence Organization). The Russian model includes elements of the intelligence services and "external" contract activity, principally the Internet Research Agency (IRA) also known as Glavset.

Iranian Cyber Operations

Iran's growing expertise and willingness to conduct aggressive cyber operations make it a major threat to national security. Their opportunistic approach to cyber attacks makes critical U.S. infrastructure susceptible targets for Iran, particularly when Tehran believes that it must demonstrate it can push back against the U.S. in other domains. Recent attacks against Israeli targets show that Iran is far more willing than before to target countries with stronger capabilities.

With the United States vowing to "have Israel's back" in response to the current conflict in which Iran-backed Hamas massacred Israeli civilians and taken hostages, the U.S. needs to look closely at Iranian Cyber capabilities being deployed. Previously Iranian hackers exfiltrated data from U.S. universities and targeted U.S. industrial control systems with the prospect of causing physical damage. Iranian Advanced Persistent Threat (APT) groups have also engaged in disinformation campaigns and supply chain attacks, and have destroyed data through wiper attacks, sometimes disguised as ransomware.

Reducing Financial Sector Cyber Risk and Cost

The rapid transition to the digital, connected world makes the national infrastructure and the financial sector subject to catastrophic attack and failure. While this digital revolution was taking place, the realm of cyber "attack" and threat also evolved from bored students hackers to major criminal organizations and hostile foreign intelligence services and militaries.

This change in the nature and level of threat was not appreciated as worldwide use of the cyber ecosystem exploded. In part, this failure of understanding can be attributed to the speed in which this threat evolved, and in part to the relatively glacial pace in which the federal bureaucracy adapts to changing missions and direction of resources to meet such rapidly evolving threats.

Meeting this challenge and enhancing the security of this infrastructure requires changes in the national approach to organization and decision-making, improved threat detection and analysis, enhanced resilience, hardening of the cyber infrastructure, and preparation for future cyber disasters. An integrated approach to the problem has several key elements:

- Ongoing collection data such as email, code artifacts, communications, and postings from hostile nations (such as China and Russia) that may contain malicious code to be used in activities such as espionage and cyber warfare.
- Use of AI tools applied to the database to identify both malicious code as well as specific individuals who are "contributors" of potentially malicious code.
- Use of graph database tools showing links between malicious code and code contributors.
- Technical experts involved in the development of offensive cyber tools that understand this software.
- Regional experts with native fluency in Chinese, Russian (and possibly Farsi) as well as expertise in cyber operations to examine the supporting infrastructure in hostile nations.

The U.S. currently lacks the ability to detect and counter hostile code development. While some agencies may collect some similar data, they lack the AI tools to evaluate the data on an ongoing basis and do not have programs in place to develop them. This means the United State lacks an effective means to prevent or mitigate a digital Pearl Harbor.

Enhancing Resilience and Hardening of the Cyber Infrastructure

For several decades computer scientists have looked at modernizing the Internet and enhancing the resilience of the existing infrastructure. Even though the Internet has gone through a period of explosive growth worldwide that was unimagined at the outset it continues to operate with many of the technologies and protocols developed in the 1960s. Resolving this aspect of the problem lies in the hands of several worldwide bodies and outside the control of any U.S agency or intelligence service.

Software ranging from operating systems to a myriad of applications remain vulnerable to hostile cyber attack. Operating systems and applications using open source elements, including especially the widely used Linux operating system, are vulnerable to "contributors" submitting patches and other changes that can introduce malware. An essential part of the mission is to monitor these contributions and their contributors with specialized AI tools for this task. Most operating systems and applications are vulnerable and need to be monitored for vulnerabilities, with specific plans made for the event they are attacked.

Development of Resilient Computer Code

To support the development of a more resilient cyber infrastructure the nation needs to provide far greater resources for software engineering institutes and similar institutions to develop much needed standards and data formats for documents and other digital media. Now there are virtually no common or accepted standards for either data formats or code development, rendering the U.S. cyber ecosystem more vulnerable to hostile attack. This should be a matter of major concern for both the Executive branch and Congress.

Meeting this challenge needs to remain a core mission for DARPA, the military services, and the Intelligence Community. Such agencies and offices have the internal structure and program management capability to execute this technical mission and need adequate resources to achieve this objective and become operational on a sufficiently large scale.

Longer-Term Measures to Prepare for a Cyber Disaster

Although no major or devastating cyber attack on the financial sector has taken place, the threat remains and will continue to grow. Longer-term measures need to be taken in order to prepare the country for the possibility of a "Digital Pearl Harbor" and respond effectively in the event it happens. Some of the most important measures include:

- Organizational changes which further empower agencies having the legal, technical, and management capability to execute an effective cybersecurity program.
- Development of a supporting infrastructure that includes an adequately funded external base of national laboratories, FFRDCs, universities, and commercial firms.
- AI technologies that meet emerging threats to critical sectors to ensure that the U.S. maintains an advantage and the ability to provide the needed resilience.
- Realistic "stress testing" and "war gaming" of potential high-end cyber attacks.
- Resilient data formats and standards for code development.
- Congressional support for funding and oversight of essential activities.

1. Introduction

Cybersecurity has been a growing concern as government, industry, and academia have all recognized the growing threats but have largely failed to adequately address the most serious aspects of the problem. Within the financial sector, as well as other critical sectors, efforts to date have largely focused on "low hanging fruit" and not on the capabilities being developed by potential adversaries or the prospect for a major cyber attack on the Internet as well as the specific systems and technologies supporting the nation's financial operations.

Enough is known about the subject and the dangers to take more comprehensive steps than ever before to enhance cybersecurity. For more than two decades, the nation has been largely complacent about the risk of a major cyber attack. The idea that the commercial sector and private industry would solve many of these problems has proved invalid, and the touted "public-private partnership" approach likely will not be totally effective.

The present context requires that the nation monitor the development of hostile cyber capabilities and the creation of malicious code as well as the institutions and individuals involved. Reporting cyber attacks as "incidents" after they take place is no substitute for understanding how a devastating cyber attack might be conducted and managed if not prevented. The results from a major incident may be no less than catastrophic, with recovery a national nightmare.

The Digital Revolution and the Internet

As a result of the digital revolution, digital data has largely replaced analog files and other media. Most of the world's communications and information technology systems have become part of the "connected world," dependent on the Internet network infrastructure. Neither government nor the private sector anticipated the speed of this technology revolution and the challenges it would pose.[1]

At the outset in the 1960s ARPA, as it was then known, initiated experiments in new technology for network optimization, with no sense that this work would evolve into the largest media revolution in history. Neither DARPA nor anybody else anticipated the extent and speed of the technological transformation wrought by the Internet and advances in computer science. Nobody at the time saw the vulnerability and security challenges that have evolved exponentially.

Communications and information technology across all sectors have become reliant on this modern infrastructure. These benefits have also provided a new venue for crime, espionage, and warfare. While the nation's financial sector rapidly adopted the new infrastructure for operations of all kinds, it has become increasingly vulnerable to a wide range of threats and malicious activities.

[1] See Abraham Wagner, Thomas Garwin, Nicholas Rostow, Sophia d'Antoine and David Aitel, *DARPA Cybersecurity Planning: Technologies for Keeping the Nation Safe* (Los Angeles: Center for Advanced Studies on Terrorism, 2018).

Over the last 30 years the Internet and reliant technologies have become increasingly central to all aspects of life. At the dawn of the Internet age there was a sophisticated view that large businesses – especially the large money center banks – would in their own interest somehow force internet technology suppliers to provide inherently secure products. Spurred by the prospect of liability and enforced by insurers and regulators, all businesses would adopt these products and related best practices, and of course governments would too.[2] Active vigilance is required today by multiple system operators, software vendors, and even users to protect systems against attack.

It may have been a historical accident that the first decade of the rise of Internet-reliant business technologies occurred at the same time as the United States, in the post-Cold War period, saw non-state actors and technologically backward regional powers with limited geographical reach as the primary threats. While the 1990's saw significant U.S. Government attention to cyber threats to critical infrastructure and worked, and the last 20 years saw increased attention following the 9/11 attacks.

The public sector and to a major extent the U.S. government have both been slow to recognize the increased danger from the cyber capabilities of China and Russia in the context of dramatically increased antagonism with these autocratic countries that have again become vociferous critics of the rules-based international geopolitical, monetary, finance, and trading regimes that the U.S. has championed since World War II. In addition to these two potential adversaries can be added Iran and North Korea in what can be termed "the cyber axis of evil."

The Expanding Attack Surface

Critical sectors all became reliant on networked digital services that provided a new venue for crime, espionage, and warfare. A new infrastructure was adopted for operations of all kinds and became highly vulnerable to a wide range of cyber threats, with an evolving attack surface. Hacking, vulnerability, resilience, and cybersecurity are now matters of great concern. Unlike kinetic warfare, cyber threats differ in terms of attribution as well as scale – ranging from the annoying and inconvenient to cyberwarfare that could disable much of the nation's operations in catastrophic cyber-attacks.

The Internet today is a far different place than it was in the early years, and it is time assess the threat landscape and the range of actions that can enhance the resilience of the financial sector. Few transactions now are done with cash and virtually all records are electronic, using systems that are all network-connected.

A major attack by a hostile state actor could result in the loss of service to key elements of the financial sector would seriously cripple and likely halt the economy.[3] Safeguards built for the

[2] The recent attack that disabled much of MGM's operations in Las Vegas shows again that this optimistic view was a mirage.

[3] See Nicholas Rostow and Abraham Wagner, *Digital Pearl Harbor: Responses to the Growing Threat* (New York: Margin Research, September 2023). See also, John Ratcliffe and Abraham Wagner, "U.S. Needs New 'Manhattan Project' to Avoid Cyber Catastrophe," *Newsweek* (May 18, 2022) and Tom O'Conner, Naveed Jamali and Fred Guterl, "Will Putin's Hackers Launch a Cyber Pearl Harbor—and a Shooting War? *Newsweek* (June 18, 2021). The global cyber attack by the WannaCry malware program, which infected over 230,000 computers in 150 countries in

"brick and mortar" world have little or no relevance to transactions in the cyberlandscape. Along with the transition to digital only, the speed of many market transactions is now measured in seconds, or micro-seconds – not days.

The attack surface continues to evolve and expand. Issues related to hacking, vulnerability, resilience, and cybersecurity are now matters of great concern as the threat landscape has continued to expand dramatically. Cyber threats differ from kinetic warfare greatly in terms of attribution as well as the scale of meaningful attacks – ranging from the annoying and inconvenient to cyberwarfare that could disable much of the nation's operations in catastrophic cyber-attacks. The ability to live with the lower range of the continuum tends to obscure the need to deal seriously with the possibility of large coordinated attacks by patient hostile actors exploiting novel as well as established vulnerabilities

.The Internet today is a far different place than it was in the early years of its development, and it is time to rethink the threat landscape facing this critical sector and the range of actions that can be taken to enhance the resilience of the financial sector.[4] The communications and IT systems utilized in the sector are all network-connected. A loss of service would largely halt the economy as it now exists.

An effective approach to the problem involves working with the Treasury Department and other federal agencies, as well as a public-private partnership with key actors within the financial sector. The present analysis includes threat modeling and threat analysis covering the broadest range of developed and emerging threats and vulnerabilities. Based on this approach, conducted in partnership with key financial sector partners, an effective research, development, and organizational strategy could be designed to address major vulnerabilities.

Both Russia and China have interests in a continuum of cyber operations against the U.S., from familiar sorts of espionage targeting government and commercial secrets, including military and technological capabilities, plans, and designs; through operations enabling blackmail against Americans and sowing distrust of the U.S. government and civil strife and election of politicians friendly to autocrats; to disruptions and threats of disruptions aimed at slowing or stopping arms transfers, military deployments, or mobilization of conventional forces to counter Chinese or Russian invasions of neighboring countries – a "Digital Pearl Harbor."[5]

May 2017, is a cautionary example of what can happen. See European Union Agency for Cybersecurity (ENISA), *WannaCry Ransomware Outburst* (May 15, 2017).

[4] Clearly some potential hostile actors are planning and investing heavily in such capabilities. With respect to China see Dave Aitel, Sophia d'Antoine, Winnona DeSombre, Isabella Garcia-Camargo, Ian Roos, Nicholas Rostow, Jonathan Smith, Alison Strongwater, Abraham Wagner, and JD Work, *China's Cyber Operations: The Rising Threat to American Security* (Margin Research, 2022), and *China's Cyber Power and Military-Civil Fusion* (New York: Margin Research, February 2023). With respect to Russia see Dave Aitel, Sophia d'Antoine, Thomas Garwin, Ian Roos, Nicholas Rostow, Justin Sherman, and Abraham Wagner, *Russia's Cyber Operations: A Threat to American National Security* (New York: Margin Research, 2023).

[5] See Nicholas Rostow and Abraham Wagner, *Digital Pearl Harbor: Responses to the Growing Threat* (New York: Margin Research, September 2023). See also, Ratcliffe and Wagner, *op. cit.* and Tom O'Conner, Naveed Jamali and Fred Guterl, "Will Putin's Hackers Launch a Cyber Pearl Harbor—and a Shooting War? *Newsweek* (June 18, 2021). The global cyber attack by the WannaCry malware program, which infected over 230,000 computers in 150

Establishing international norms to limit government-supported cyber intrusion activities has been difficult because of the lawless behavior of certain governments and is not helped by common understanding that the U.S. Government has engaged in parts of the cyber intrusion continuum abroad and that significant secret hacking tools have been criminally released from U.S. Intelligence agencies.[6]

The cyber intrusion continuum and the difficulty of attribution allows hostile cyber forces to practice their skills and even establish persistent presence in American IT systems and those of our allies without much sanction at the same time as it lulls some into thinking that cyber intrusions and attacks are a manageable cost of doing business. In contrast to kinetic attacks, cyber attacks do not require exquisite tactical coordination or courageous warriors willing to risk their lives; there has been little downside to failed attempts at penetrating critical systems. It is also the case that cyber attacks, unlike kinetic ones, can scale from lower level espionage or covert attacks to major ones that can disable the financial sector – giving the attacker a far wider set of options.

Though Russia and China have the greatest combinations of strategic motives and technical skill and sizeable cyber offensive forces, the nature of the technologies and Russia's and China's interest in obscuring their own operations and also Russia's weakness and reliance on Iran and North Korea for weapons for use in Ukraine means that threats from regional powers such as Iran and North Korea cannot be ignored. Nor can disruptions from criminal elements associated with hostile government cyber activities.[7]

The cyber threat may seem manageable until access to or even hidden persistent presence in U.S. networks and zero-day exploits are used in a coordinated and strategic fashion by a hostile power to produce effects at an intensity and scale completely beyond anything experienced so far – or even just to provide a coercive demonstration of such effects.

Deterring Cyber Attack

Deterrence has long been fundamental to U.S. national security strategy. Potential adversaries, including nation states and non-state actors such as terrorist groups and criminal enterprises, have been deterred from attacking the U.S. and allied nations because of the unacceptable costs from retaliation, both conventional and nuclear. As the range of possible attacks now includes cyberwarfare, policy and strategy for deterrence is needed to incorporate this new domain of espionage and warfare. This remains an essential pillar of the nation's overall cybersecurity strategy.[8]

---

countries in May 2017, is a cautionary example of what can happen. See European Union Agency for Cybersecurity (ENISA), *WannaCry Ransomware Outburst*, (May 15, 2017).

[6] See for example https://www.wired.com/2017/03/wikileaks-cia-hacks-dump/.

[7] See Aitel, et al, *DARPA Cybersecurity Planning, op. cit.*

[8] See Abraham Wagner, *Deterring Cyberattack* (June 2021). Previously the December 2017 *National Security Strategy,* the January 2018 *National Defense Strategy* and the February 2018 *Nuclear Policy Review* set forth and approach to meeting these challenges at a time of changing threats and technologies employed by potential adversaries including those posed by cyberattacks. This strategy placed a high priority on meeting cybersecurity goals that support deterrence and responding effectively to cyberattack and information warfare.

The digital revolution created a world where digital data has replaced analog files and other antiquated media while most of the world's communications and information technology systems have become part of a "connected world" dependent on the Internet network infrastructure. Neither government nor the private sector anticipated the speed of this technology revolution and the challenges it would pose.

Cybersecurity issues related to hacking, vulnerability, denial of service, and information warfare are now matters of great concern involving not only vital national security operations but power, finance, and other critical sectors. The concepts of defense and national security have needed to adapt, and now incorporate cyberwarfare as a major conflict domain.

Major sectors, including national security, power, finance, and others quickly adopted these technologies and became highly dependent on the commercial infrastructure enabling them. Needed investments in technology to secure this infrastructure were not made. Now vulnerabilities are more broadly recognized, and the government more committed to addressing them.

Deterrence against Russian and Chinese cyber attacks, already made difficult by the continuum of cyber intrusion, is of course made more difficult by their survivable nuclear forces which will inevitably make U.S. leaders cautious in escalating a conflict with these powers, and also by the likely difficulty in convincing global audiences that U.S. attribution of cyber attacks to one or the other of these governments is correct.

These problems with deterrence mean that protection of U.S. critical systems from attack, and rapid restoration in the event of attack damage, are crucial. Certainly China's great technical capacity and dominant position in many supply chains means that managing the U.S.-China trade and political relations and trade and technology interdependencies will also likely be key to avoiding disastrous cycles of geopolitical escalation.[9]

Technology Trends

At the same time as the geopolitical character of cyber threats has changed, few in the public or the government as a whole fully understand how changes in technology and business complicated efforts to secure computer networks and data systems against sophisticated and sustained attacks of the sort have that China and Russia are equipped to execute, and how foreseeable (though not certain) technological developments could pose even more problems.[10]

Technology trends already affecting cybersecurity operations include:

---

[9] On China's preparations for cyber operations both outside and within the context of a hot war. See Dave Aitel, Sophia d'Antoine, Winnona DeSombre, Isabella Garcia-Camargo, Ian Roos, Nicholas Rostow, Jonathan Smith, Alison Strongwater, Abraham Wagner, and J.D. Work, *China's Cyber Operations: The Rising Threat to American Security* (Margin Research, 2022), and *China's Cyber Power and Military-Civil Fusion* (New York: Margin Research, February 2023). With respect to Russia see Dave Aitel, Sophia d'Antoine, Thomas Garwin, Ian Roos, Nicholas Rostow, Justin Sherman, and Abraham Wagner, *Russia's Cyber Operations: A Threat to American National Security* (New York: Margin Research, 2023).

[10] Added to these two can be Iran and North Korea (DPRK), although these two actors have more limited technical capabilities but are moving rapidly to enhance their capabilities and active operations.

- Increased requirements for systems to be accessible and even operated and maintained from remote locations – organizations do not have 'perimeters' anymore;
- Increased overall complexity of software systems;
- Increased reliance on open-source software as (sometimes hidden) elements in the overall software ecosystem, which can have bugs and vulnerabilities introduced by malicious actors with no locus of trust to anyone;
- "DevOps" rapid software improvement cycles and increased diversity of hardware and software platforms requiring multiple versions to coexist in software-in-use and requiring complex continual maintenance providing additional paths for vulnerabilities to be created and routes for malicious access;
- Increased complexity and automation of software supply chains and maintenance by multiple third parties, creating additional vectors for rapid and widespread introduction of cyber vulnerabilities;
- Increasing reliance on software execution "in the cloud" – on servers not owned or controlled by the either the corporation whose data is being processed or even a IT contractor to that corporation, using operating systems that may not be known outside the cloud provider.

These recent technological trends could soon be joined by other complications, including:

- Use of Large Language Models (LLMs) to write and patch code – creating another possible vector for malicious code insertion for example by "indirect prompt injection" or the accidental proliferation of code with common errors or vulnerabilities;
- Use of chatbots and deepfakes to enhance phishing and other social engineering attacks; and
- Quantum computing or other breakthroughs allowing the defeat of cryptography-based security schemes.

While some of these technologies can be applied along with others to reduce certain sorts of cyber vulnerabilities, they all combine with the shortage of well-trained and loyal IT professionals to exacerbate the one constant in cyber vulnerability – the importance of "social engineering" in providing unauthorized high-level access to computer networks and systems that can be exploited in ruinous ways.

Vulnerability of the Financial Sector

The present analysis examines the financial services sector of the U.S. economy – its importance to the economy and our ability to project military force, its dependencies on international partners, its reliance on computer networks and technologies, its resulting cyber vulnerabilities, its efforts to mitigate them, and technological improvements that could further increase their security. A companion analysis reviews the legal and regulatory framework under

which banks and insurance companies and related business operate and how regulations should be changed to increase cyber security in the financial services sector.[11]

The U.S. financial services sector, in cooperation with the U.S. government, and also international financial standard setting bodies, have devoted substantial attention and resources to cyber security. Many view the situation as an arms race between cyber attack capabilities and cyber defense/resilience – with the emphasis on ability to restore operations with the assumption that breaches will occur. Cyber attacks are viewed as the number one risk to large banks by Chief Risk Officers of Global Systemically Important Banks. This shows the seriousness of the threat and also that it is being addressed.[12] Nevertheless the tools at hand may not be adequate in the face of Russian and Chinese capabilities in the context of a geopolitical crisis or hot war.

Even against normal "peacetime" cyber threats, the year-to-year data suggests that little progress has been made in five years in reducing corporate losses or reducing the time to recognize an intrusion or fix the problem after an intrusion is recognized.[13] Moreover the U.S. financial sector also has unique aspects that heighten vulnerability and the economic and national-security consequences of this vulnerability, in contrast both to other sectors of the American economy and to the financial services sectors of other developed countries. No other large sector of the U.S. economy is as dependent on distributed and external information access for its basic operations.

Today money and investments are essentially digital and need to be exposed to external access to be useful. No other sector is as necessary to day-to-day operations of businesses and households. None depends as much on the confidence of market participants that would be damaged by public knowledge of a successful cyber attack. Compared both to other sectors of the U.S. economy and to the financial sectors of other countries, the U.S. banking sector is amazingly decentralized and diverse, ranging from community banks to money center commercial banks to specialized investment institutions, yet all are closely interconnected and internationally exposed.

Ensuring the cyber security of the U.S. financial services sector into the future will require a combination of heightened vigilance in the near term and the application of existing and novel technologies to increase the inherent safety, active defense, and resilience of financial IT systems in the longer term, combined with a robust national effort to understand and counter cyber intrusion and attack threats and blunt them closer to their sources. It is clear from over 30 years of history that financial institutions and IT companies will not achieve this result on their own.

Enough is known about the subject and the dangers to take more comprehensive steps than ever before to enhance cybersecurity. For more than two decades, the United States has been

---

[11] Gaelin Bernstein, Nicholas Rostow, Alison Strongwater and Abraham Wagner, *Defending the Financial Sector Against Cyber Threats: Legal and Regulatory Environment* (New York: Margin Research, December 2023).

[12] January 2023 EY-IIF survey (Institute of International Finance, Ernst and Young). https://www.iif.com/portals/0/Files/content/32370132_ey-iif_global_bank_risk_management_survey_2022_final.pdf

[13] *IBM Security Cost of a Data Breach Report 2023.* This annual report is based on a global multi-industry survey of CIOs who have had data breaches. https://www.ibm.com/reports/data-breach

complacent about this risk of major cyber attack. The idea that the commercial sector and private industry would solve many cyber problems has proved invalid, and even the touted "public-private partnership" likely will not be totally effective.

The present cyber context requires that the nation monitor the development of hostile cyber capabilities and the creation of malicious code as well as the institutions and individuals involved. Reporting cyber attacks as "incidents" after they take place is no substitute for understanding how a devastating cyber attack might be conducted and managed if not prevented.

The country needs to take a far more disciplined approach to dealing with the threat, including detection of malicious computer code development so that the attack surface is made less vulnerable, and critical systems are made more resilient. We need to ensure that cyber vulnerabilities of financial services and other critical sectors do not contribute to hostile powers' believing that they would easily prevail in a confrontation or hot war against U.S. interests.

2. Financial Sector Functions, Composition and Importance

Major Financial Institutions

By all accounts the financial sector is a critical component of the nation's infrastructure. It includes commercial banks and related activities, securities exchanged, brokers and dealers, as well as providers of the key financial systems and services that support these functions. The sector comprises some $108 trillion in assets and faces a wide variety of cybersecurity-related risks.[14]

On the one hand there is the potential for monetary gains from criminal activities undertaken against firms and others in the sector. Apart from criminal and monetary gains there is a more serious concern coming from hostile actors that may seek to harm the nation with economic disruptions that can be accomplished without a kinetic attack. This greatly increases the financial services sector's attractiveness as a target for malicious actors.[15] The key risks include:

- An increase in reliance on and frequency of data and software interactions with outside information technology service providers and supply chain partners
- Growth in sophistication of malware, and other malicious software
- Increased interconnectivity through the use of networks, cloud service providers, and mobile applications

The nation's major institutions and firms are all totally reliant on networked systems. Within the financial sector the major components are considered below, all of which are highly vulnerable to cyber attacks of all kinds. As the world has moved to largely cashless operations these institutions maintain all records online; their customers use their web sites and applications for most transactions; card systems; as well as automated clearing house (ACH) systems employed are all vulnerable to hacking and other form of malicious cyber operations.

Trading on the NYSE, ASE, NSADAQ, CBOT, the OTC market, and other market exchanges for securities and their derivatives are entirely electronic and Internet-dependent. The SWIFT system for international transfers as well as some title operations and brokerages have been electronic for years and also vulnerable to malicious cyber activities.

[14] See Government Accountability Office, *Critical Infrastructure Protection: Treasury Needs to Improve Tracking of Financial Sector Cybersecurity Risk Mitigation Efforts* (September 2020).

[15] High-profile breaches at commercial entities, such as Equifax, have heightened concerns that data are not being adequately protected. See Government Accountability Office, *Data Protection: Actions Taken by Equifax and Federal Agencies in Response to the 2017 Breach* (Washington: August 30, 2018).

Cashless Payment Systems

The payments systems run by the banking system and related financial services and government organs like the Federal Reserve System have a role in the economy analogous to the circulatory system of the human body. On a somewhat longer time average scale than executing transactions and providing liquidity, financial services institutions and related government organizations also are essential to providing loans. They also serve as custodians and official record keepers for the vast majority of the assets and wealth of Americans. The banking system is also important to macroeconomic stability through its role in regulating the supply of money. Thus the importance of proper functioning of financial services businesses is out of proportion to the under 8% of GDP that Financial Services and Insurance accounted for in 2022.[16]

Sixty years ago, transactions were dominated by cash and paper checks. Securities orders were settled after three business days, and checks were accepted for eventual settlement. Americans could go for days or weeks without executing a bank transaction. Asset ownership including investments, deeds and mortgages and other loans were also recorded on paper. It was also a time when users had control over their own data and had some degree of choice in the matter.

To the extent that they existed, data processing systems were entirely secure except perhaps from insider manipulation or theft. They operated in batch mode and the runs could always be repeated with the same input if a problem were discovered. Networks did not exist and few computers were connected in any way.

Today the situation is very different where all operations are completely reliant on real-time transaction approval and processing for things like purchasing gas and food.[17] It is almost impossible to exaggerate how necessary an operating financial services sector – and especially banking and payment systems—is to day-to-day life, the economy, political stability, and even force-projection and mobilization in the event of crisis or war.

Compared to most other countries, the U.S. financial sector appears quite diverse and even decentralized. Even after decades of consolidations there are still over 4,000 commercial and almost 600 savings banks insured by the Federal Deposit Insurance Corporation (FDIC), of which nearly 2,700 commercial banks and 300 savings banks are supervised by the FDIC. There are also nearly 3,000 federally-chartered and nearly 1,800 state-chartered NCUA-insured credit unions.[18] Some of these banks focus on local communities while others (like Silicon Valley Bank) focus on serving specific industries and attracting the business of specific types of wealthy individuals.

---

[16] Bureau of Economic Analysis, https://www.bea.gov/sites/default/files/2023-06/gdp1q23_3rd.pdf, Table 14.

[17] Unlike an earlier era, virtually all retail transactions – no matter how small – are done with either credit cards or iPhone applications and similar technologies. The biggest exception now may be illicit drug deals although there is evidence that even some of these transactions are now being done with apps.

[18] https://www.fdic.gov/analysis/quarterly-banking-profile/statistics-at-a-glance/2023jun/industry.pdf and https://ncua.gov/newsroom/press-release/2023/credit-union-assets-shares-and-deposits-grow-fourth-quarter#:~:text=The%20number%20of%20federally%20insured,%2C%20state%2Dchartered%20credit%20unions.

The Impact of Crypto-Currency

Increasingly, crypto-currency exchanges and block chains can also serve as stores of value and transactions processing alternatives to banks; but these still play a small role in the overall economy. Unfortunately, decentralization and diversity are not necessarily a strength in terms of resistance and resilience in the face of cyber attack, as many of the smaller community banks have less than state-of-the-art defenses and restoration plans and capabilities in place, many are dependent on a few large financial and information processing vendors, and all are dependent on correspondent and other payment processing relationships with a few credit card networks and a few money center banks

Moreover, compared to most businesses, banking is uniquely dependent on confidence. As anyone who has seen *It's a Wonderful Life* knows, depository institutions lend multiples of their deposit base and are subject to bank runs if consumers and counter-parties become concerned that access to their money may be compromised. Thus, as a result, problems in one bank may cause problems for otherwise unaffected banks. This is the opposite of the situation with other goods and services providers, where an inability of one business to serve customers is likely to generate increased total demand via precautionary buying.

The seemingly diversified structure of the American banking system may in fact produce a situation where interdependencies between less- and more- vulnerable institutions add vulnerability overall rather than making the system more robust. It would be cold comfort to, for example, deploying soldiers from military bases that tend to be located in rural areas, that the rest of the country was unaffected if they were unable to access funds from their accounts or receive their direct-deposit paychecks owing to a targeted cyber attack on the possibly less-sophisticated smaller banks and credit unions serving their area.

Other Parts of the Financial Infrastructure

In addition to banks, several other sorts of companies are crucial to the modern functioning payments system in the U.S. While bank regulators have the authority to inspect and regulate the companies that provide processing services to banks, it is not clear that they are effectively reaching all the organizations that provide intermediary services not purchased directly by banks. Banks have traditionally relied on third-party financial infrastructure organizations, sometimes called financial utilities, for check clearing, electronic funds transfers and wires which are automated clearing house functions.

Banks now generally purchase online bill pay services to provide to their customers and issue Master Card and VISA branded debit and credit cards. American Express and Discover issue cards directly and process payments through their own networks. Point-of-Sale and online payment processors and terminal providers such as Square (Cube), Shopify, Toast, Zettle by PayPal, Stripe, and Clover provide different bundles of services to merchants, though most merchants sign up for only one of these services.

These newer payment systems generally process payments through the platforms of money center banks and either appear to the banks as a single (aggregating) merchant of record or as an independent sales organization with a contractual relationship to the money center bank. Many

may sell data about sales patterns as an additional revenue source. PayPal provides another interface that allows payments from bank accounts and credit cards, and mobile payment apps such as Venmo, CashApp and the bank-sponsored Zelle also provide transfer and payment functions, and may facilitate the growth of payment via crypto currency in the future.

In theory at least, all of these companies could be subject to various sorts of harmful data breaches and cyber attacks. In the short term merchants may be reliant on one of these but they are close substitutes so changeover is possible in theory, though with the possibility of substantial losses and changeover costs. This value of diversity is, however, offset by the likelihood that the various companies may share common IT providers and substrates and thus common vulnerabilities.

Consumer Credit Bureaus, such as Experian, Equifax, and Transunion, are important for such events as new loans and increases to credit lines.[19] Business reporting agencies such as Dun and Bradstreet perform similar functions for business. The fact that these services are used less frequently than payment systems and the fact that merchants and banks can use more than one makes them somewhat less crucial in the event of a short-term service interruption.

Another set of companies that provide many bank-like services are retail investment houses like Fidelity Investments, Vanguard, TD Ameritrade, and others. These companies provide not just stock and bond and mutual fund trading and custodianship but also check writing for cash management accounts, generally through a commercial bank but the investment company provides the interface, credit cards, and other bank-like services.

Money Center Banks have their own unique characteristics – primarily in that their lending is financed by money market borrowing – including internationally, more than by deposits, and in their greater international operations. At one point in time there was a stark distinction between these commercial banks and investment banks, although this separation is not currently as clear as in earlier times.

In addition to stabilizing the banking system against bank runs through the creation of the Federal Deposit Insurance Corporation, the Banking (Glass Steagall) Act of 1933 separated commercial and investment banking, largely preventing commercial banks from dealing in or underwriting securities and preventing investment banks from retail banking. While this distinction was erased by legislation in 1999, it still marks the financial institutions in place today.

Key Role of the Federal Reserve System

The entire range of banking and quasi-banking services discussed above are all offered within the context of an overall monetary and payment system run by the Federal Reserve System, including the Federal Reserve Board and the twelve Federal Reserve Banks, that operates as the Central Bank of the United States. As such, the Federal Reserve System:

[19] Equifax was the victim of one of the largest data exfiltrations in history. See Government Accountability Office, *Data Protection: Actions Taken by Equifax and Federal Agencies in Response to the 2017 Breach*, (Washington: August 30, 2018).

- Conducts the nation's monetary policy to promote maximum employment and stable prices in the U.S. economy;
- Promotes the stability of the financial system and seeks to minimize and contain systemic risks through active monitoring and engagement in the U.S. and abroad;
- Promotes the safety and soundness of individual financial institutions and monitors their impact on the financial system as a whole;
- Fosters payment and settlement system safety and efficiency through services to the banking industry and U.S. government that facilitate U.S.-dollar transactions and payments; and
- Promotes consumer protection and community development through consumer-focused supervision and examination, research and analysis of emerging consumer issues and trends, community economic development activities, and administration of consumer laws and regulations.[20]

The Federal Reserve does not have exclusive responsibility for many of these functions within the United States. In addition to the Federal Reserve, other government and international entities have key roles affecting financial sector operations:

- The Treasury Department has a huge presence in financial markets through its operations to fund the Federal government debt, and affects corporate incentives through Internal Revenue Service interpretations of tax law. Obligations of the United States Government and its agencies provide a major share of global liquid interest-bearing assets.
- The Treasury Department, through the FDIC and the Office of the Controller of the Currency, plays key roles in supervising retail banks and thrift institutions.
- Investment activities are regulated by the Securities Exchange Commission and the Commodity Futures Trading Commission. These two agencies have been odds over who should regulate crypto currencies.
- State bank and insurance regulators also have authorities that affect corporations within their jurisdiction.

Internationally, U.S. Financial institutions operate with counterparties regulated by other countries' central banks and regulatory systems, process payments through international payment networks such as SWIFT, participate in international credit card and other systems. Central banks operate within the context of multilateral financial institutions, such as the Bank for International Settlements, the International Monetary Fund, the World Bank, regional development banks, and

[20] *Money and Payments: The U.S. Dollar in the Age of Digital Transformation* (January 2022). https://www.federalreserve.gov/publications/files/money-and-payments-20220120.pdf

the G-7 Financial Stability Board, and challengers to some of these institutions sponsored by China.

Finance as it relates to international trade may also be affected by World Trade Organization rulings. Both internationally and domestically, it can be challenging for all these government entities to move rapidly and in a coordinated fashion to do what is needed to provide appropriate regulation in the face of innovation.

With the rise of new payment modalities such as instant transfers, record keeping systems including various cryptographic block chain systems and distributed ledgers, asset classes (cryptocurrencies and tokenized assts and NFTs) and exchange mediums (cryptocurrencies), many of the basic concepts underlying these money supply, payment system, and regulatory functions are being re-learned because distinctions that had become forgotten by settled day-to-day operations are becoming newly important in the context of innovation. Thinking about future cybersecurity of the financial sector, five directions of potential change are worth mentioning, with brief descriptions of how they might affect the current financial system and its stability, especially in the context of imperfect cybersecurity.

Instant Payments and Settlement

Worldwide people now make more than two billion digital payments a day.[21] Until the experimental FedNow system went live in the summer of 2023, there were no actual instant payments in the United States, except where cash physically changed hands.[22] Where one orders an automated clearing house (ACH) or wire transfer from their bank, it normally takes hours or even a day or two for the transaction to be completed and settled, and so it could be backed out in case of error. When a payment is made by credit or debit card, this creates an immediate claim against the users account that is later settled between the merchant and the payment network and ultimately removed from the users account. For the most part these complications are a distinction without much of a difference, but they could become important in the event of a successful cyber attack.

In a normal monetary environment with high certainty of stability of the institutions involved and low inflation, it doesn't matter to ordinary people that the only money that is actually "central bank money" is cash and balances held by banks and the U.S. government in the Federal Reserve System; the rest is claims against bank-issued money that can be converted at par to central bank money. There are, however, circumstances, including a cyber breach, where the distinction could become important to the parties involved – for example if a chain of instant transfers based on fraudulent access could lead to the money disappearing without a trace, rather than as would have been the case in the old system, stopped at one or two points removed in a multi-day process.

Cryptocurrency as a Medium of Exchange

It is fair to say that, at least for now, the bloom is off the private cryptocurrency rose. Companies like Meta that were moving ahead to issue their own cryptocurrency shelved the

[21] BIS *Red Book Statistics* cited in https://www.bis.org/publ/arpdf/ar2022e3.htm.

[22] https://www.federalreserve.gov/newsevents/pressreleases/other20230720a.htm.

projects. Few have followed El Salvador in making Bitcoin legal tender. Bitcoin had a split personality since its inception – as a speculative financial asset and a currency. It is already fairly clear that stability of value is important for a medium of exchange but not for investment.

The huge amount of energy consumption required by the Bitcoin "proof of work" mining model has given all of crypto a bad name, however unfairly. Crypto "Stablecoins" established to maintain parity with the U.S. dollar or a basket of commodities have not yet come into wide use and some have experienced catastrophic failure.[23] There is a strong argument that stablecoins are a solution in search of a problem.[24]

While the model of an immutable transaction record is intrinsically attractive, the realities of implementation are far from transparent or readily understandable. For example, Ethereum, the second most popular cryptocurrency, and one that does not require mining but instead is based on a "proof of stake" model, requires an-ever changing assortment of participants to achieve "consensus" on whether transaction records are to be accepted onto the block chain, raising questions of whether the block chain operations could be disrupted or compromised by a determined effort to affect the computers doing the voting.[25]

Blockchain promoters claim that they can be secure, decentralized, and scalable to handle the very large number of transactions required of normal currencies. In 2017 increased blockchain usage led to congestion and dramatically increased transaction fees. While there are solutions to this issue, they tend to increase complexity and cost. The Bitcoin network experienced a similar episode in 2023, suggesting that the problem has not been solved. None of these networks has demonstrated scalability that would be required to take over a substantial slice of U.S. payments.[26]

Researchers have talked about the "Blockchain Scalability Trilemma," suggesting that blockchains can have two of the three desired attributes, but not all – they can be secure and decentralized but not scalable, or secure and scalable but not decentralized, or scalable and decentralized, but not secure.[27] Though computational solutions such as "sharding" have been advanced, the increase the complexity and opacity of the platform's operation, as well as the cost.[28]

Cryptocurrencies in isolation are unusable, except perhaps for speculation. In a world with multiple cryptocurrencies coexisting with each other and with traditional money, cryptocurrency owners need to exchange their crypto currencies for money, for goods and services and for other crypto currencies. To make crypto currencies usable, crypto exchanges and crypto-friendly banks grew up. Problems have emerged with both of these types of institutions which have yet to be fully resolved.

---

[23] For example the TerraUSD which lost all its value in May 2022. https://www.bis.org/publ/arpdf/ar2022e3.htm.

[24] https://cepr.org/voxeu/columns/stablecoin-paradox.

[25] https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/.

[26] *Ibid.*

[27] Adrian Tobias, "Cryptocurrencies and Decentralized Finance," Bank for International Settlements 21st Annual Conference, June 24, 2022. https://www.bis.org/publ/work1061.pdf#page=74.

[28] *Ibid.*

The exchanges rather than the blockchains themselves have been the main successful target for cyber theft.[29] Unlike banks and investment companies, these exchanges are not regulated in the public interest. The failure of the FTX exchange and the criminal convictions of its leadership revealed that that particular exchange was in at least part a criminal enterprise with a political operation to protect it from government attention.[30]

Cryptocurrency owners would also like to be able to easily exchange their tokens for actual money and goods and services in the real economy. Economists have noticed that this would have the effect of essentially enlarging the money supply and making the implementation of monetary policy more difficult.

More immediately, it leads to the danger of problems in the unsupervised crypto world being transmitted to the regulated banking system. Signature Bank was the first FDIC-insured bank to establish a payment system that tied crypto to the banking system with near 24 hour availability. Because of its friendliness to depositors who were crypto enthusiasts. it lost deposits in the wake of the FTX failure and was subject to a run and closed by regulators once Silicon Valley Bank went under.[31]

Finally, while it has proved possible for law enforcement to track criminal uses of crypto-currency for example to receive ransomware payments, the degree to which cryptocurrency can allow bad actors to evade reporting requirements, counter-terrorist restrictions, international sanctions, and use crypto for crime remains a serious concern that will affect policy.[32]

Tokenized Assets and Self-executing "Smart Contracts"

Current practice is for financial institutions including mortgage lenders and insurance companies to deal with physical objects by reference to documentation. There is some potential for these assets to be "tokenized"-- essentially title would be assigned within a digital ledger rather than in the legal system. A digital ledger can also include self-executing future transactions conditional on a combination of states of the world coming into existence that are cognized by the digital ledger. Obvious new vulnerabilities are possible, to the extent that the terms of these smart contracts or the tokenized assets themselves could be changed by cyber attack. It is also the case that some assets are natively digital (NFTs) and could of course be traded via a blockchain.

---

[29] Chainalysis, *The 2023 Crypto Crime Report*. (February 2023), https://go.chainalysis.com/rs/503-FAP-074/images/Crypto_Crime_Report_2023.pdf. Chainalysis reports $3.8 Billion stolen by hacking in 2022, including $1.65 Billion by North Korea. Distributed Finance protocols including cross-chain bridges and smart contracts were the source of the vast majority of these stolen funds. Here some $386 million was lost not to traditional hacking but by fooling crypto exchanges into misjudging market conditions and prices for various cryptocurrencies.

[30] See Darreonna Davis, "What Happened To FTX? The Crypto Exchange Fund's Collapse Explained," Forbes (June 2, 2023).

[31] https://apnews.com/article/signature-bank-fdic-new-york-svb-40c361918e2bc9c20d7b19b683b01f65.

[32] Department of the Treasury, *Illicit Finance Risk Assessment of Decentralized Finance*, (April 2023). https://home.treasury.gov/system/files/136/DeFi-Risk-Full-Review.pdf .

Central Bank Digital Currency and Universal Ledgers

Many of the touted benefits of private cyber currencies might be secured, and the manifest current problems avoided, if central banks were to take responsibility for issuing "Central Bank Digital Currency" (CBDC) based in a blockchain or "Universal Digital Ledger." The blockchain could in theory be either centralized or decentralized, or hybrid with it maintained privately for ordinary purposes but backed up authoritatively by the central bank in case of cyberattack or other problems with the decentralized system.

Immediate payments and transfers could be accommodated. The digital currency could be interest-bearing or not, or both. It could be issued directly to individuals or just to banks.[33] If just to banks, they and other financial institutions could provide their customers with digital currency accounts pegged to the CBDC but they would not be legal claims on the Federal Reserve system.[34] In the extreme CBDC record keeping could extend to tokenized assets and smart contracts. Many would be repelled by what would seem to be a government takeover of what might have previously appeared to be private functions.[35]

As reflected in analysis by the Federal Reserve Board, the Bank for International Settlements, and others, the degree to which CBDC might affect the stability and performance of the financial sector depend dramatically on the design details, even assuming the system works perfectly. The provision of direct access to central bank accounts would likely reduce the incentive for individuals and firms to keep money in banks, especially if the CBDC accounts paid interest. This would perforce reduce lending by banks and thrifts.

Even non-interest-bearing accounts at the Fed would dramatically increase the amount of central-bank issued money in the economy, potentially causing problems for monetary policy. On the other hand such accounts could dramatically reduce the cost of international transfers. This can happen only if other countries were accommodating and could increase access to funds in the event of natural disasters or cyber attacks on private institutions, again assuming the system worked.

Over 114 central banks are currently considering issuing a CBDC. The Bahama Sand Dollar has been in circulation since 2021. More importantly, in China the digital renminbi (e-

[33] Under current law, the Federal Reserve System is not allowed to take deposits from or issue currency directly to individuals, so legislation would be required for individual CBDC wallets or accounts in the United States.

[34] Board of Governors of the Federal Reserve System, *Money and Payments: The U.S. Dollar in the Age of Digital Transformation,* (January 2022) https://www.federalreserve.gov/publications/files/money-and-payments-20220120.pdf.

[35] BIS 2020 *Annual Economic Report* (June 2020) Chapter III. Central banks and payments in digital era https://www.bis.org/publ/arpdf/ar2020e3.htm and BIS annual report 2022 III. The future monetary system vhttps://www.bis.org/publ/arpdf/ar2022e3.htm, February 9 2022 The Future of Money IMF managing director speech at Atlantic Council https://www.imf.org/en/News/Articles/2022/02/09/sp020922-the-future-of-money-gearing-up-for-central-bank-digital-currency, https://www.imf.org/en/Publications/fintech-notes/Issues/2022/02/07/Behind-the-Scenes-of-Central-Bank-Digital-Currency-512174, https://www.cfr.org/backgrounder/cryptocurrencies-digital-dollars-and-future-money.

CNY) has more than on hundred million individual users and billions of yuan in transactions.[36] The system uses private sector banks (or what passes for private sector in China) to provide accounts to customers. There are, however, limits on holdings and the government is leveraging the system for even more insight into the transactions of users.[37]

So far these CBDC systems are operating alongside conventional banking systems. If they ever largely took over a monetary system there would be many effects both on financial stability, financial privacy, and of course cyber security – depending on the details. These details include policy issues in designing the CBDC, technical and privacy issues in designing its accounts and accounting ledger system, whether it includes programmable smart contracts and tokenized assets, and regulatory issues to the extent that part of the system is administered by private sector institutions.

[36] https://www.imf.org/en/News/Articles/2022/02/09/sp020922-the-future-of-money-gearing-up-for-central-bank-digital-currency.

[37] https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-central-bank-digital-currency-cbdc

3. Trends in Financial Sector Cyber Risk

Cybersecurity as a New National Security Challenge

The bank robber Willie Sutton is famously reported to have answered, when asked why he robbed banks, "Because that's where the money is." Banks also have data worth stealing, including significant amounts of confidential client and counter-party information. At least for now, the primary method for the theft of money from financial institutions is with digital tools, including phishing or to otherwise access account credentials from customers or from servers outside the direct control of banks.

There is considerable reason for concern, as banks and other financial institutions complete the digitization of all their records and systems, that much more calamitous results could emerge from successful cyber attacks, even where the objective is not the theft of money or data. Financial institutions have essentially all the IT systems and functions of any business, with the possible exception of some specialized process control systems that are characteristic of manufacturing or the electric grid. Their "inventory" is now essentially all data – deposits and obligations are information records – no longer physical entities.

While there have been many publicly reported data breaches and even successful ransomware attacks on established U.S. financial institutions, there have been no spectacularly successful cyber attacks that have drained bank accounts for large numbers of customers, prevented access to funds, or destroyed account records.[38] This has not been true elsewhere, for example in the UK. More commonly, customer information and credit card and automated teller machine card information is stolen that can later be used for theft. Reported losses of $40 million CNA Financial paid in ransom or a few million dollars here and there in ATM theft using pirated data are small compared to financial industry profits and have not inconvenienced customers.

By contrast, large crypto-currency thefts totaling in the hundreds of millions of dollars have been reported just in the last few years, suggesting that established regulated financial institutions have greater cyber defenses and resilience than the more immature crypto ecosystem.[39] In the last

[38] Publicly available information on U.S. banking disruptions due to cyber attacks should now be more complete, though many aspects of breaches may remain obscure. Since May 1, 2022, U..S. banking organizations have been required to notify their Federal regulators and customers of computer security incidents affecting them or their banking services providers that has caused or is reasonably likely to cause a disruption or degradation of service to banking customers lasting more than four hours. https://www.federalregister.gov/documents/2021/11/23/2021-25510/computer-security-incident-notification-requirements-for-banking-organizations-and-their-bank.

[39] The Carnegie Endowment for International Peace provides a useful listing of major cyber attacks on financial institutions that can be filtered by the country whose institutions were attacked, the purpose of the attack, whether it was state sponsored, etc. https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline#click-hide. A Congressional Research Service report summarizes Justice Department Press Releases describing cyber-attacks attributed to state actors and foreign criminals from 2012-2022.

twenty years, cybersecurity has become a major industry with North American revenue estimated to have been $68-billion in 2022, staffed largely by graduates of university cyber security programs.[40] A 2020 survey showed that Banks and other financial institutions spend heavily on cybersecurity – typically around 11% of IT spending, 0.5% of revenue, and $2,700 per FTE.[41]

Nevertheless, around the globe both government leaders and financial institution insiders are as concerned as they have ever been about the vulnerabilities of financial institutions to cyber attack. A global survey of Financial Institution Chief Risk Officers (CROs) released at the beginning of 2023 suggests that despite billions invested to safeguard core systems and protect vital data assets, CROs consider cyber the top inherent threat and the one most likely to result in a crisis or major operational disruption.

Even when they perceive their own internal systems as largely secure, CROs see potential amplifications and concentration of cyber risk lurking everywhere — within geopolitical turbulence, ecosystem strategies and the vast networks of partners, suppliers and vendors on which banks increasingly rely. The interconnectedness of those networks — and the integrated technology that underpins the entire global financial system — represents a massive attack surface and a huge perimeter to secure. Because bad actors are relentless in seeking vulnerabilities and because successful attacks are so lucrative, it's worth asking if cyber and other threats to resilience will ever recede very far from the top of CRO agendas.[42]

The data strongly suggests that survey respondents representing Global Systemically Important Banks (G-SIBs) and European institutions especially linked their increasing concern about cybersecurity with the potential for increased attacks from Russia and China as a result of the deteriorated geopolitical situation. The concern is not limited to large banks. In the U.S., 74% of executives representing almost 250 community banks view cyber security as among their top three concerns, more than any other risk.[43]

Despite all the effort and focus, as the survey data on the level of concern indicates, there is no feeling that the cybersecurity problem has been solved, and data show no meaningful positive trends in time to discover a network incursion or data breach, or time to fix the problem once

---

https://crsreports.congress.gov/product/pdf/R/R46974. For crypto theft see https://go.chainalysis.com/rs/503-FAP-074/images/Crypto_Crime_Report_2023.pdf.

[40] https://www.fortunebusinessinsights.com/industry-reports/cyber-security-market-101165

[41] https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/risk/Cybersecurity.pdf Banks and Financial Utilities spent somewhat larger amounts than insurance companies and non-bank consumer financial service firms. Spending as a percent of revenue is similar to that reported for hospitals, yet hospitals more than banks seem to have been plagued by successful ransomware attacks. Data from HHS study- https://405d.hhs.gov/Documents/405d-hospital-resiliency-analysis.pdf Possibly the difference relates to the longer-standing focus on cybersecurity at banks and financial institutions.

[42] *Seeking Stability Within Volatility: How Interdependent Risks Put CROs at the Heart of the Banking Business*. 12th annual EY/IIF global bank risk management survey. https://www.iif.com/portals/0/Files/content/32370132_ey-iif_global_bank_risk_management_survey_2022_final.pdf.

[43] https://www.aba.com/-/media/documents/reference-and-guides/2023-banking-and-risk-compliance-report.pdf

discovered.[44] Rather there a continuing arms race between defense methods and attacker competence, with many attackers protected behind the shield of non-cooperative foreign governments even if their identities are discovered. CEOs and corporate board as well as cyber professional are aware that financial firms are perhaps an undiscovered or unpatched vulnerability or an employee mistake away from a sizeable loss of data and even a substantial business interruption.

Moreover, the results would be much worse if not for concerted efforts by a small number of companies, in some cases aided by the U.S. Intelligence Community, to monitor new types of attacks and track the changing tactics, procedures, and exploits of specific attackers, so that appropriate vendors and target companies can provide and apply patches and other countermeasures rapidly.[45]

Recognition that cyber security defense will always be challenging because of the role of the human factor and social engineering, and that it is especially challenging now because of recent and future changes in the financial services IT ecosystem, has led to adding a major focus on cyber resilience – the ability to operate through and rapidly recover from even a successful attack -- to the earlier focus on static defense – firewalls, cybersecurity education and hygiene, access limits based on authentication. More recently there has been a focus on dynamic defense, such as intensive monitoring of network traffic and software execution to detect signatures of attack and compromise, and rapid response as new sorts of attacks are recognized.[46]

Trends in Technology and the Financial Services Ecosystem

Trends in technology and the financial services ecosystem are making cybersecurity staffs feel that they have less control than in the past and somewhat less confident that technologies are available to redress the balance. In looking through these details, it's worth bearing in mind an overall definition of cyber risk that recognizes the potential social costs of a successful attack, the dynamic relationship between actual technological vulnerability and attacker competence, and the extent to which cyber resilience capabilities can avoid the potential costs of even a successful attack. In a notional equation,

---

[44] According to the IBM/Ponemon Institute Cost of a Data Breach survey of CISOs, in 2023 it took 204 days on average to recognize that a breach had occurred, and 73 days to contain the breach once noticed, an increase from 2018 when it took 197 days to identify an additional 69 days to contain a data breach. https://www.ibm.com/downloads/cas/E3G5JMBP.

[45] See recent reports from CrowdStrike, Akamai, Verizon, and Contrast Security. https://go.crowdstrike.com/rs/281-OBQ-266/images/CrowdStrike2023GlobalThreatReport.pdf , https://www.akamai.com/site/en/documents/state-of-the-internet/soti-enemy-at-the-gates-analyzing-attacks-on-financial-services.pdf , https://www.akamai.com/resources/research-paper/akamai-ransomware-threat-report, and www.verizon.com/dbir [annual Data Breach Investigations Report], and https://www.contrastsecurity.com/hubfs/Cyber%20Bank%20Heists%20Report%202023.pdf?hsLang=en.

[46] The emphasis on cyber resilience is notable not only among cybersecurity consultants and vendors but even among regulators and international standard setting bodies. See the analysis from the Financial Stability Institute of the Bank for International Settlements on *Banks' cyber security - a second generation of regulatory approaches*. https://www.bis.org/fsi/publ/insights50.htm. See also https://www.cisco.com/c/en/us/solutions/hybrid-work/what-is-cyber-resilience.html, https://www.pnnl.gov/explainer-articles/cyber-resilience, https://www.cyberresilience.com, https://www.ibm.com/topics/cyber-resilience.

Cyber Risk = (Vulnerability * Attacker Competence and Focus) * Potential Social Cost / Cyber Resilience, where vulnerability is the actual exposed vulnerability net of defensive security measures.

In a more detailed model, it is possible to focus on multiple separate vulnerabilities down to a quite micro level, and then sum this equation for an overall risk, though in actuality because of the need for complex attacks exploiting multiple vulnerabilities in sequence to succeed against a defended system, one would need to take such complex attacks into account by replacing the first term by the product of individual vulnerabilities and attacker competence and focus on that vulnerability for all of the elements required to execute the complex attack that would create the damage.

In addition to cyber risk, in arriving at an overall national policy one must also address cybersecurity cost – the day-to-day expense of cyber defense and cyber resilience technologies, consultants, and operations because alternative investment strategies may achieve the same level of cyber risk at very different costs.

Increasing Risk from Complex Financial Sector Technology Ecosystems

Cybersecurity is a continuing problem because modern information processing systems and network infrastructure are made up of networked general purpose computer processors running complex heterogeneous stored programs made up of common operating systems, various software utilities, and application programs that themselves are congeries of software modules, some of dubious provenance and security. To do their jobs, these systems are interconnected and connected to the internet, and they are designed to execute programs and to be updated under remote control. Software vulnerabilities or unauthorized access via stolen or credentials open these processors to being remotely reprogrammed for nefarious purposes, including launching attacks on other computers that can be reached over a network, corrupting, erasing, encrypting or exfiltrating data, and even causing physical damage through control interfaces.

In complex attacks, initial access, often achieved by logging on with credentials secured by phishing or guessing, is expanded by various means, including opening backdoors for further access, stealing additional logon credentials, escalating access to administrator or superuser level that allows more data access and more extreme processor reprogramming. The attacker may establish a persistent presence within the target network, storing and forwarding information at will for later exfiltration, and inserting programs into various processors that can be triggered later –conceivably months or years later – as part of a culminating attack, all the while using various methods such as "fileless" attacks to disguise their presence in the network from scans designed to detect viruses, malware, and other attack signatures.

Even simple attacks based on indiscriminate scanning and broadcast can cause extensive damage, extending even to perhaps unintended targets. Software vulnerability to cyber attack has been exacerbated by a variety of trends in the software and hardware landscape that all make software more complex and more heterogenous, running on more diverse sorts of processors that are more interconnected and more open to the outside internet, with more routes for the insertion of accidentally or deliberately created vulnerabilities. These trends include more:

- Requirements for systems to be accessible and even operated and maintained from remote locations – organizations do not have 'perimeters' anymore;
- Complexity and automation of software supply chains and maintenance by multiple third parties, creating additional vectors for rapid and widespread introduction of cyber vulnerabilities;
- Reliance on software-as-a-service, distributed processing, and execution "in the cloud";
- Dependence on complex software supply chains including open-source software as (sometimes hidden) elements in the overall software ecosystem, which can have bugs and vulnerabilities introduced by malicious actors with no locus of trust to anyone; and
- Layers of software installation and processes running to achieve defense against and resilience after a cyber attack.

Remote Access

Financial Institutions have been offering online and mobile account access for years, but the COVID-19 pandemic accelerated the number of employees working remotely, including IT maintainers and developers and cyber-security specialists needing high-level access. Remote access by personnel working home was also needed by employees of third-party software specialists and developers.

The result was dramatically increased requirements for authentication of privileged access to the IT systems of Financial Institutions and their IT and cyber-security contractors.[47] Desktop sharing, allowing complete remote access to a desktop, often provided to allow remote IT diagnosis and repair, created an initial attack vector that has been estimated to have been recently used in about a third of ransomware attacks, second only to email phishing as the mechanism of access.[48]

Third Parties, Complex Software Supply Chains, and Frequent Updates

Corporate Chief Information Officers are acutely aware that they do not have control over risks to their networks, data, and system posed by "third parties."[49] Third parties include payments systems and intermediaries that access the a financial firms IT systems through data interfaces or Application Program Interfaces (APIs), providers of Software as a Service (SaaS) that interface intimately with the companies data and computers, and software developers and maintainers performing outsourced IT functions.

They are also concerned about risks from "Shadow IT" software installed on computers without approval of corporate IT departments. According to some definitions, firms may have

[47] https://www.imf.org/-/media/Files/Publications/covid19-special-notes/en-special-series-on-covid-19-cybersecurity-of-remote-work-during-pandemic.ashx

[48] Verizon *2023 Data Breach Investigation Report*, Figure 31.

[49] *Ibid.*

thousands of third parties that could be the source of vulnerabilities or attacks.[50] Many of these partners provide software that is incorporated in the core of the firm's operations or maintain software and thus have intimate access. Both afford additional attack vectors as well.

Many of the most serious recent cyber vulnerabilities have emerged from third party software. The NotPetya attack by Russia on Ukraine that wreaked havoc in 2017 destroyed the global shipping firm Maersk's ability to operate because of a back door Russian military hackers had inserted in software widely used for filing Ukrainian taxes, a copy of which was installed on one Maersk computer in Odessa.[51]

A more typical example is more recent. In early 2023, a ransomware gang called CLOP (or CL0P) began exploiting a previously unknown vulnerability of Progress Software's MOVEit enterprise file transfer tool that provided access to various SQL database server applications. In June, Progress Software rapidly issued a patch once the vulnerability was discovered. In the meantime, however, CLOP stole data from government, public, and business organizations worldwide – affecting at least 2000 organizations and data on 62 million people.[52]

Typical software products are made up of agglomerations of hundreds or thousands of modules of unclear provenance, leaving end-user IT departments an imperfect view of their vulnerabilities. The complexity of the installed software base, the frequent discovery of new bugs, the need to inter-operate with a constantly changing hardware and software ecosystem, and the competitive drive to add new features and improvements in the DevOps rapid deployment cycle all contribute to software providers pushing frequent updates that can become the vector for attacks.

In 2021, for example, hackers believed to be associated with the Russian Military Intelligence Service (SVR) were able to penetrate the Texas network software provider SolarWinds and add malicious code to its update of the Orion network monitoring appliance. A routine update infected U.S. companies including Microsoft, CISCO, and Intel and Federal Departments and Agencies including Treasury, Justice, Defense and the Cybersecurity and Infrastructure Security Agency (CISA) in the Department of Homeland Security (DHS). The attack installed back doors and was largely undetected for nine months.[53]

---

[50] See also https://cdn2.hubspot.net/hubfs/2378677/Ponemon2020_Final.pdf.

[51] https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/ Only the discovery of one domain controller in Ghana that was accidentally powered down for the duration of the attack as a result of a power failure allowed Maersk to restore its system.

[52] IBM was among the companies whose servers – hosting data from health providers and others –were breached. A massive amount of data was also stolen from life insurance companies including Prudential. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-158a https://www.theverge.com/23892245/moveit-cyberattacks-clop-ransomware-government-business.

[53] https://www.npr.org/2021/04/16/985439655/a-worst-nightmare-cyberattack-the-untold-story-of-the-solarwinds-hack.

Distributed and Cloud Processing

Increasingly, data and transactions are processed by multiple computers receiving instructions from one another, all under software control. This of course requires mechanisms for one computer to instruct another what to execute, and these mechanisms can be exploited to execute malicious code if it can only be injected into the system at an appropriate place. Unfortunately there all too many locations in current code environments were such instructions to execute malicious code can be inserted and hidden.

The cloud – with code loaded and executed on servers owned by such companies as Microsoft, Google, IBM, Oracle, or Amazon, introduces its own complications for security, as employees of the cloud may be able to access data from users and the links between the user and the cloud may be susceptible to exploitation as well. Cyber risks from cloud computing are among the top concerns of financial sector IT managers.[54]

Open-Source Software

Software is written in high-level languages that are understandable to human programmers as well as computer compilers or interpreters that convert the "source code" to binary instructions that can be executed by computer processors. The source code for an open-source program is publicly available for inspection. Distributed open-source development projects welcome code contributions, corrections, and improvements from anyone which means that bugs and vulnerabilities can be introduced by anyone, either by accident or maliciously. While substantial resources are devoted to limiting this risk in various ways, such efforts are uneven.[55]

Open-source software is now ubiquitous. The Linux operating system is the most prominent example of a large open-source project. Variants of Linux are predominant globally in a wide range of product types – including supercomputers, web servers, mobile phones, network appliances such as routers, and the popular raspberry pi single board computer that is increasingly used for embedded computing in small-batch commercial products.

The only major exceptions are personal computers where Microsoft Windows and Apple MacOS make up the majority; tablets, where Apple is about even with the Linux based Android, and mainframe computers, where Linux makes up a large minority of installed systems but IBM's

[54] A notorious example involved an Amazon Web Services employee who used her knowledge of cloud server vulnerabilities to steal personal information of over 100 million people from more than two dozen corporate entities, including most prominently Capital One Bank. https://www.justice.gov/usao-wdwa/pr/former-hacker-sentenced-stealing-computer-power-mine-cryptocurrency-and-stealing.

[55] Whitepaper "Open Source Security Census: Open Source Software Projects Needing Security Investments" by the Institute for Defense Analyses and the Linux Foundation and *Census II of Free and Open Source Software — Application Libraries* (March 2022). The Linux Foundation and The Laboratory for Innovation Science at Harvard Frank Nagle, Harvard Business School James Dana, Harvard Business School Jennifer Hoffman, Laboratory for Innovation Science at Harvard Steven Randazzo, Laboratory for Innovation Science at Harvard Yanuo Zhou, Harvard Business School https://8112310.fs1.hubspotusercontent-na1.net/hubfs/8112310/LF%20Research/Harvard%20Census%20II%20of%20Free%20and%20Open%20Source%20Software%20-%20Report.pdf.

z/OS is still dominant in IBM installations.[56] Most web servers run open source web servers such as nginx or Apache on top of the Linux operating system.[57]

Since most proprietary software products and services incorporate open source code to perform basic functions, there is potential for vulnerabilities in open-source software packages to affect proprietary products as well. Major enterprise software providers, including Microsoft, Oracle, and IBM, all now incorporate open source packages in their products and contribute to open source development.[58] Smaller and more specialized software providers are even more likely to incorporate open-source software utilities in their products, and apply fewer resources to ensure that any vulnerabilities introduced by open-source packages are contained.

Problems associated with open-source code are compounded by heterogeneous and complex software supply chains. An enterprise may purchase install open-source software packages directly, in which case it is responsible for updating frequently and applying the most recent patches, or it may purchase eEnterprise Open-Source Software from a vendor that adds an additional layer of vigilance, or it may rely on outsourced IT administrators to keep its software up to date, or a SaaS provider to do the same. The purchase or download of additional software packages that contain modules that were created some time in the past may also no longer be maintained, while employees may download and install additional software packages without approval of the IT department.

Examples of major cyber vulnerabilities in enterprise systems caused by open source code are legion. Notably the recent log4shell exploit combined open source and supply chain vulnerabilities. Log4j is an open-source java logging utility that is widely deployed in enterprises as part of a variety of software packages. The logging function means that the module injests and logs many events. To exploit the vulnerability all a hacker needs to do is to get the module to read and log a specific character string, in almost any context. A typical enterprise computer might have many instances of Log4j buried in different applications and utilities.

Since HTTP traffic for website access and interchange among computers via websockets is often logged, it's relatively easy to gain access. Then the host computer can be taken over or effectively destroyed. Ironically the problem could have been much worse if companies kept their

---

[56] https://en.wikipedia.org/wiki/Usage_share_of_operating_systems, It is notable that both CISCO's and Juniper Networks' proprietary operating systems for their routers are now based on Linux. Such appliances used to have custom code that was essentially not hackable but the need to be able to add new features and rapidly deploy new products caused the switch to Linux.

[57] Most web servers run open source web servers such as nginx or Apache on top of the Linux operating system. Although proprietary web server software from Cloudflare and Microsoft is increasing in market share, Cloudflare acknowledges that its offerings are built on open source server (nginx) and database platformshttps://www.netcraft.com/blog/may-2023-web-server-survey/ , https://blog.cloudflare.com/open-source-two-way-street/. Other commonly used open-source projects include Python (now the most commonly used high-level language), C++, Microsoft .NET, javascript, java, ruby, and R.

[58] See https://www.ibm.com/opensource/story/ , https://opensource.microsoft.com/ , https://en.wikipedia.org/wiki/Microsoft_and_open_source, https://www.oracle.com/a/otn/docs/oracles-commitment-to-open-source.pdf. See also Ibrahim Haddad, *A Guide to Enterprise Open Source: Developing and Executing Open Source Software Strategy* (May 2022). s://8112310.fs1.hubspotusercontent-na1.net/hubfs/8112310/LF%20Research/LF%20Research%20Guide%20to%20Enterprise%20Open%20Source.pdf.

software up to date, because the vulnerability was introduced in Log4j Version 2 and after the vulnerability was found it was discovered that version 1 was still ten times more prevalent, despite having known unpatched vulnerabilities and having been declared end-of-life and no longer supported since 2015.[59] Obviously this is an indication that the problem of obsolete software modules is pervasive and serious.

Old open-source software embedded in deployed modules that are no longer being actively maintained can expose systems to penetration and worse. For example the Microsoft Threat Intelligence Center reported in November 2022 that the open-source Boa web-server, a project abandoned in 2005, still ships as part of a variety of Internet of Things (IoT) devices and popular software development kits (SDKs). A vulnerability in the Boa web server was exploited as the entry point for attacks on the Indian electric grid by an attacker suspected to be associated with the Chinese government from 2020 through 2022.[60]

Cyber Defense

Outsourced cyber defense and software and systems used for cyber defense and cyber resilience purposes can also be vectors for attack. For them to work, anti-virus and anti-malware systems must have access to most if not all file structures and even the operating system. Such defenses must be updated frequently as new vulnerabilities, exploits, and malware become known. So all the characteristics are in place for system compromise by this route.

Virtual Private Networks (VPNs) and cloud back-up are other services that are intended to provide security but can he weak points. VPNs promise secure transit but unencrypted data sent over them are visible to the VPN provider. Vulnerabilities in software packages that transfer encrypted data, such as OpenSSL, can expose even encrypted data to reading by the VPN provider or other transit points on the internet.

Back-up and remote storage services provide routes for exfiltration and generally complete access to the data being backed up. As more and more AI-based supervision systems are developed, they may also degrade system performance by stopping execution of benign processes if the AI is imperfect or subject to malicious injections.

Current and Emerging Technologies and Techniques to Reduce Cyber Risk

In discussing various measures to enhance cyber security and resilience, it is worth keeping in mind the "CIA triad" of data protection – it's important to protect the Confidentiality, Integrity, and Availability of data and data processing systems. Confidentiality means that data can be accessed only by those authorized to do so. Integrity has to do with the data continuing to be in its proper state (complete and accurate, not corrupted or falsified). Data and system availability

[59] https://www.wired.com/story/log4j-log4shell-one-year-later/ and https://www.wired.com/story/log4j-log4shell-vulnerability-ransomware-second-wave/. On prevalence of various versions of Log4J, See FOSS Census II, *op cit.*

[60] https://www.microsoft.com/en-us/security/blog/2022/11/22/vulnerable-sdk-components-lead-to-supply-chain-risks-in-iot-and-ot-environments/.

includes timely access to and manipulation of existing data and appropriate timely recording of new data and newly processed information.[61]

Attacks may target and affect one, two, or all three of these dimensions. Compromises of confidentiality can result in irremediable damage as response may be equivalent to closing the barn door when the horse is gone. Damage to data integrity, when recognized, can in theory be repaired if the information required to do so is available, though permanent damage can be done via actions taken based on inaccurate data before the damage is recognized or the repair is made. Denial of service attacks are inherently reversible in terms of the information technology system but even they can have irreversible consequences in the real world, for example if critical transaction services are interrupted or if crucial transient data isn't acquired.

Focusing on seemingly ubiquitous vulnerabilities shouldn't obscure the significant efforts and resources that have been devoted to defending against attack, for operating through attacks, and for rapidly restoring service following a successful attack. Many of the successful attacks mentioned above had to be executed in very clever ways to circumvent defensive and quality assurance measures.

Broadly speaking, current and emerging efforts can usefully be divided into several categories and subcategories:

- *Static Defense*
  - Control Access Based on Identity
  - Restrict Inward and Outbound Traffic and Software Installed on Endpoints
  - Ensure Correct Routing and Confidential Data Exchange
  - Track Data Provenance
  - Obscure and Extend the Perimeter
- *Reducing Vulnerabilities in Deployed Systems*
  - Train Staff and Enforce IT Hygiene
  - Respond to Announced Vulnerabilities: Contain and Patch
  - Focus on Software Supply Chains
- *Resilient Operation Under Attack*
  - Centralize network security information and automate response
  - Use Systemic Threat Intelligence
  - Exercise and Test
  - From Back-up and Recover to Robust Operations

The categories broadly reflect the evolution of approaches to cyber security. Within each category and subcategory, there has been a progression over time, largely in response to related

[61] See for example https://www.nccoe.nist.gov/publication/1800-26/VolA/index.html. So far as is known, the "CIA Triad" has no connection with either the Central Intelligence Agency or the Culinary Institute of America.

threats becoming more sophisticated. Some technological and topological trends are common across categories, including increasing use of artificial intelligence techniques beyond rule application, including machine learning, and roles for service providers that are positioned to monitor a wide swath of public and private network activity to discern new threats and vulnerabilities.

Static Defense

Static Perimeter Defense is the cyber equivalent of a border fence, although "firewall" is the usual word used for devices that enforce policies to separate the inside from the outside. It's an effort to keep the data processing system and internal network free of outside agents, and prevent unauthorized exfiltration of data, while maintaining needed bi-directional interchange with the outside for authorized data flows and transactions.

In practice, perimeter defense techniques may be applied in multiple layers of a network topology, partly because networks are assembled from pieces (including "endpoint" devices such as desktop computers) designed to operate on their own and partly because networks typically have multiple segments where connection to the outside is possible, each with their own router on which firewall software is typically installed.

Ultimately the multiple required accesses in a modern enterprise largely invalidated the metaphor of a perimeter at the same time as the prospect of massive Distributed Denial of Service (DDOS) attacks made a perimeter with a small number of portals a liability. In the more distributed architecture typical of modern enterprise computer networks, the preferred metaphor now is "Zero Trust Network Access," in which no user or computer process receives access to a computing resource or data just by virtue of being inside the fence – instead access is allowed based on policy rules and frequent external authentication of the user and user processes.

Control Access Based on Identity and Permissions

Even before the Internet, time-shared and other remotely accessible computer systems required usernames and passwords to allow access and to regulate levels of access and service based on user identity. Now people continue to use user names and passwords for access to personal computers and various accounts on networked computers.

Many of the security issues now bedeviling networked computers were initially recognized in the time-shared environment, where storage and memory might include data owned by various users and so there was a need to be sure that processes launched by one user didn't access data owned by another, and that users couldn't use hacks to upgrade their access privileges to the administrator level.[62]

[62] Time shared systems were vulnerable to unauthorized use of various system functions and also memory overflow exploits. There is an argument that today's operating systems have security challenges partly because they are all descended from or influenced by Unix. Unix left out many of the security features from the earlier Multics OS whose development had been initiated by MIT, GE, and Bell Labs. See Alexander Klizhentaz, "In Search of the Perfect Access Control System," *Teleport Blog* (March 4, 2021). https://goteleport.com/blog/access-controls/ Also https://en.wikipedia.org/wiki/Multics.

From the point of view of security, the ideal would be a system that checked before each new process was launched or any process accessed additional data that this particular usage is within the rights of an absolutely verified user and that the access made sense in terms of the role of that user. For a variety of historical reasons having to do with the early environment in which networks and the Internet and existing desktop operating systems were designed, current systems do not achieve this level of control, and so there are a lot of band-aids being applied instead.

At the most basic level, the Internet lacks inherent mechanisms for establishing user identity. The combination of a user name and password is still the most common mechanism for authenticating access but runs into well-known problems. Users too often use simple-to-remember passwords or don't change the default, often using the same password for more than one system, so that any compromise of a password list endangers other systems as well. Logon credentials are now solicited and brokered on the dark web, and in extreme cases may be provided by disloyal employees.

Urged to change their passwords frequently and use "strong," hard-to-remember passwords, users store passwords in insecure places or forget them entirely, requiring mechanisms for password reset that open additional routes for credential compromise either via social engineering attacks on the help desk or access to an email account or cell phone text messaging system used for secondary authentication.[63]

Answers to secondary challenge questions are notoriously easy to guess based on social media or other information about individuals that may be widely available. Ironically, challenge questions that allow password resets are often much less "strong" than the passwords themselves, and access to an email account means compromise of all accounts for that user protected by only passwords resettable by access to the email account or by resettable passwords and two-factor authentication by email. Recognizing that a password is only as secure as the password-reset system, Microsoft and others now offer password-less logins based on authentication from a mobile device. The mobile device then becomes a route for network compromise.

Multi-factor authentication schemes layered on top of password-controlled access enhance security in the normal course of events by providing notice of attempted logins and requiring an action from a separate device. They also open additional routes for credential compromise as they have to have provisions such as one-time codes or other authentication routes for when the secondary authentication device is unavailable.

Cryptographic tokens, whether hardware or software, and even biometric signatures are only as secure as the control over the token and information by the user. Even user-oriented password managers such as Dashlane or LastPass, or Google's browser- and cloud- based password storage, or Apple's cloud and device based keychain could both provide access to multiple systems in the event of a device or system compromise. Instances where thieves have drugged users and gained access to their financial accounts by the facial recognition function on

[63] Research sponsored by DARPA and others for some time has demonstrated that even the best passwords provide only limited security against well developed hacks. Recently multi-factor identification has improved things somewhat but is far from the security needed.

their phones are a reminder of this.[64] It is possible that high-level access to otherwise secure systems could be gained by similar attacks on system operators and maintainers.

Another approach pushes authentication via the OAuth open authentication protocol onto some other provider with a more direct relation to the user – such as invitations to authenticate via Google, Apple, Microsoft, or Facebook. These schemes may provide information to leak customer information to web sites as well as provide unauthorized access in the event that Google, Apple, Microsoft, or Facebook develop a security issue.

Enterprises may avail themselves of third-party authentication services offered by companies specializing in this area, such as Cloudflare and Okta. These providers allow options for differentiated access to various resources (various corporate systems APIs, web servers, data, and account information, etc.) based not only on authenticated user identity but also on the specific device and IP address making the access, providing additional security and protection. We will return to these sorts of third party security providers when we discuss extending the enterprise perimeter, below.

Service from such third-party providers is only more secure to the extent that the provider of these services remains uncompromised or their systems are arranged to be somehow ignorant of the actual information required for authentication.[65] An attack on or service outage at the authentication provider can preclude access to an enterprise's networks, possibly bringing operations to a halt, even if it does not affect data confidentiality or integrity. Once again, a layer added to improve security has the potential to cause additional problems.

Looking at some 40 years of identity-based access control is that there has been a history of changes to the authentication regime to deal with the changing mechanisms for achieving unauthorized access. Each innovation tends to increase complexity and provide additional potential routes for compromise. Overall, unauthorized access via stolen, phished, or guessed logon credentials, even on systems where multi-factor authentication is required, nevertheless remains a primary initial attack route against enterprise networks and systems.[66]

Restrict Inward and Outbound Traffic and Installed Software. Firewalls, typically now implemented in software on endpoint computers and in routers or other high-speed processors at network segment boundaries as well as at portals to the internet, selectively allow or block connections and pass or block traffic based on policies designed to allow needed information to pass and to block unneeded and potentially injurious traffic.

Similarly, anti-virus and end-point restrictions block installation of software and files that can be injurious, including in many enterprises not just network interchange but hardware ports

---

[64] *Ibid.*

[65] Okta was reported to have been compromised as recently as 2022. https://thenewstack.io/the-okta-mess-is-even-worse-than-it-appears/. Its help desk has recently as September 2023 been the target of social engineering efforts to achieve MFA resets to enable use of previously secured usernames and passwords to access systems of U.S. customers. https://www.securityweek.com/okta-says-us-customers-targeted-in-sophisticated-attacks/.

[66] According to the Verizon *2023 Data Breach Investigation Report*, over 60% of breaches involved stolen credentials, phishing, or pretexting.

such as USB (formerly floppy disk) "sneakernet" connections to prevent infection and installation of unauthorized software as well as unauthorized data exfiltration.

Network firewalls have become more sophisticated over time.[67] The earliest "stateless" packet filters dropped packets if they lacked appropriate source and destination IP addresses, source and destination port numbers, and data transport protocol. These filters could not keep track of whether packets were part of an ongoing connection or not, and so left networks vulnerable to malicious traffic that was minimally similar to normal usage, for example faking the source IP address. "Stateful" firewalls were developed to track connection status to help filter out these counterfeit packets.

The next step in firewall innovation added packet inspection capabilities, allowing traffic to be discarded based its information content – for example if it contained known virus files, identified by a hash, signatures, or anomalous protocol activity. In addition to discarding harmful packets, high speed firewalls could also preferentially discard traffic intended to flood specific servers or applications in a denial-of-service attack. Logging and alert functions aided in threat diagnosis and warning.

Outgoing traffic rules could prevent enterprise users from accessing dangerous web sites or exfiltrating data either to unknown IP addresses or to known bad IP addresses, or in some cases based on the content. Over the last decade, firewalls and firewall-like technologies have extended rules-based filtering based on specific user permissions, web-application usage, and even the content of encrypted packets.

Unfortunately, hackers have developed ways around many of these defenses. Anti-virus and anti-malware scans that search for filenames and file content have been defeated by "fileless" attacks and exploits that hide code outside of the file system and hijack native utilities to do the work.[68] Encrypted payloads may evade even deep packet inspection. As the firewalls became more sophisticated in function, they needed more frequent updates, and so manufacturers changed from relatively secure hardware or custom operating system architectures to open-source operating systems, largely Linux or BSD, and even open-source router software that has been vulnerable to exploitation by hackers.

In an effort to make firewalls more responsive to and useful against rapidly evolving threats, vendors have begun to embed advanced pattern recognition and other machine learning (ML) technologies in firewall software. Because machine learning is not perfect, it's plausible that sometimes the system will over-generalize and block necessary traffic as well. Expected benefits include:

---

[67] For the history of firewalls, see https://www.juniper.net/documentation/en_US/learn-about/LA_FIrewallEvolution.pdf, https://www.paloaltonetworks.com/cyberpedia/the-evolution-of-firewalls-from-packet-filtering-to-machine-learning-powered-ngfws, and Kenneth Ingham and Stephanie Forrest, *A History and Survey of Network Firewalls*. (University of New Mexico technical report 2002). https://www.cs.unm.edu/~treport/tr/02-12/firewall.pdf

[68] See for example https://www.crowdstrike.com/cybersecurity-101/malware/fileless-malware/.

- Rapid automatic adjustment of rules to filter out malicious traffic based on similarities to previously identified threats.
- More rapid recognition and blocking of source IP addresses that have been taken over by attackers.

Finally, cloud-based firewalls or router-installed firewalls connected to specialized security vendors that actively monitor a very large number of enterprise firewalls can take advantage of rapid automated detection of new threats seen across the web and ML-based changes to rule sets automatically and rapidly deployed. It is at least theoretically possible that such immediate supervision of firewalls could become a route to allow attackers in if the vendor system were compromised.

In 2022, firewall vendors FortiGuard, SonicWall, Sophos, and Zyxel all had to release security advisories and patches, and some of these vulnerabilities had been actively exploited for months previously.[69] As with the history of authentication, firewalls have evolved as attacks have become more sophisticated, but web-based attacks still succeed.

Ensure Confidential Data Exchange

Over the last decade, web traffic has increasingly become encrypted by the SSL and TLS protocols and the use of encrypted Virtual Private Networks (VPNs) and wi-fi connections, in order to prevent data and credentials theft in transit across the internet. Advances in computer power rendered early encryption standards susceptible to unauthorized decryption. Encrypted transport can be an obstacle to perimeter defense by packet inspection, making it a matter of two steps forward, one step back. Moreover, encryption and decryption utilities are additional processing steps that can, at least in theory, be hijacked to exploit data, and vulnerabilities in both encryption utilities and VPNs have been discovered, including ones deemed to be of "high severity."[70]

Track Integrity and Provenance of Data and Information

Increasing use of "deepfake" and other technologies for creating convincing but fake documentation, as well as the discovery of hostile forces developing persistent presence in enterprise and government networks, have raised the possibility of deliberate insertion of false data and information for nefarious purposes. While back-ups and journaled information systems can help retrospectively, it would be better to have mechanisms in place to warn of tampering and to ensure that data presented is accurate.

One approach is analogous to the "parity check" or "checksum," originally applied to fallible computer memory, but now used to check file transfers, that statistically helps demonstrate

[69] 2023 VZ DBIR "Year in Review."

[70] See https://insights.sei.cmu.edu/blog/vpn-a-gateway-for-vulnerabilities/, and an account of Chinese hackers accessing Airbus and Rolls Royce technical secrets by way of their VPNs in 2019. https://media.cert.europa.eu/static/MEMO/2019/TLP-WHITE-CERT-EU-MEMO-190927-2.pdf. Also https://papers.mathyvanhoef.com/usenix2023-tunnelcrack.pdf and https://www.openssl.org/news/vulnerabilities.html.

that data hasn't been altered, though without disclosing the nature of any alteration. Such methods are not generally appropriate for hostile actors who may be able to alter data in a way that such external summary checks remain valid. Digital watermarks may be applied in a less-public fashion that then is less susceptible to faking the validation signature.[71]

Obscure and Extend the Perimeter

Massive, distributed denial-of-service attacks (DDOS) can cause problems for almost any individual enterprise's firewall and connections to the internet. To avoid this, alternate enterprise IP addresses can be hidden and traffic routed around the globe through multiple access points if needed to provide enough bandwidth to deal with the issue. Vendors such as Cloudflare provide this service, aggregating many enterprise data flows through their access points and thus decreasing the vulnerability of any one enterprise.[72]

Again, the addition of another vendor provides another potential attack vector. This trend toward an obscured and extended perimeter is occurring in the context toward cloud computing, software-as-a-service, outsourced authentication and other security services, and more complex software and maintenance supply chains, all of which are working against the very notion of an effective network perimeter.

Reducing Vulnerabilities in Deployed Systems

It is important to think about how enterprises are advised to minimize the vulnerabilities in deployed systems that remain despite the installation of the technologies discussed above. This category has a lot to do with how systems are organized and managed rather than just the application of technology.

Train Staff and Enforce IT Hygiene. Human error is a factor in the vast majority of network intrusions and data breaches. The first piece of advice generally given is to train staff well in good cyber security practices including how to recognize phishing and other "social engineering" attacks. Keeping updating software to the latest version and applying patches as soon as they are issued is a standard recommendation that is not always followed, and personnel should be discouraged from connecting unsecured devices to the enterprise network and devices and from installing un-approved software. Systems and training should enforce strong passwords and multi-factor authentication.

Respond to Announced Vulnerabilities: Contain and Patch

Software vendors have the primary responsibility for disclosing vulnerabilities that have been discovered in their software and to advise their customers on how to mitigate the vulnerabilities, sometimes through temporary measures to contain the problem followed by application of "patches" – software revisions to mitigate or eliminate the vulnerability. The ability to respond in a timely way may be limited in some contexts by the availability of trained personnel.

---

[71] DARPA has some efforts underway to develop technologies for making sure that digital documents are at least tamper-evident. See the DARPA SafeDocs program. https://www.darpa.mil/program/safe-documents.

[72] https://www.cloudflare.com/ddos/#DDoS-Page-Pricing-AS.

At times, however, the enterprise is not the direct customer of the software with the vulnerability, and owing to the complexity of software supply chains there may be known unpatched vulnerabilities lurking in packages supplied by third parties. To a greater degree than is commonly recognized, often initial containment instructions and patches limit the initial route that was found to exploit an underlying vulnerability but do not conclusively eliminate the vulnerability, requiring additional patches as attackers find new ways around the initial containment efforts and patches.[73]

Behind the announcement of vulnerabilities and how to deal with them, there is now a small amount of infrastructure that encourages publication and helps focus discussion.[74] This process helps maintain consistency and focus in understanding the vulnerability, the degree to which it is being exploited, and efforts to contain the exploitation and patch the vulnerable software. Increasingly new attack vectors and vulnerabilities are recognized not by software suppliers but by cybersecurity vendors that have installed software across thousands or even millions of networks and machines and receive current reports of unusual activity from these systems.

Focus on Software Supply Chains and Open-Source Software Including Embedded Utilities. The SolarWinds and Log4Shell vulnerabilities have dramatically heightened attention in the cybersecurity community to software supply chain attacks and the propagation of open-source software utilities throughout the deployed software base, bringing their problems with them. For now the best advice to enterprises is primarily to keep versions up to date and patch as soon as possible, but new approaches are emerging.

The sloppiness at SolarWinds, where it is rumored that access was gain by using a user name and "Password123," should not obscure the fact that even SolarWinds used sophisticated software audit tools to attempt to ensure the integrity of its software, which the attackers were able to get around only by inserting their malware just before the binary was created.[75]

Since 2018, the National Telecommunications and Information Administration (NTIA) in the Department of Commerce has been leading a multi-stakeholder process to advance the concept of the Software Bill of Materials (SBOM). There are several competing approaches to maintaining an SBOM in machine readable form, updated as needed with each new software release. At least in concept, having a complete and current SBOM would allow rapid judgments about whether a software package is likely to contain a particular vulnerability and would make clear whether the software contains modules that are no longer being maintained or have known unpatched vulnerabilities.

---

[73] See CrowdStrike *2023 Global Threat Report*, p.14.

[74] Since 1999, the Common Vulnerabilities and Exposures (CVE) program has been operated by MITRE on behalf of the Department of Homeland Security (DHS) Cybersecurity and Infrastructure Security Agency (CISA). MITRE and other CVE Numbering Authorities (CNAs) control the issuance and maintenance of CVE IDs to track disclosed vulnerabilities. The National Institute of Standards and Technology (NIST) has a team that assigns a common weakness enumeration (CWE) category and assesses the exploitability and impact of the vulnerability and includes this information in the National Vulnerability Database (NVD). https://nvd.nist.gov/general/cve-process#.

[75] See *Wired, op. cit.*

Researchers have attempted to develop vulnerability risk scores depending on a variety of variables for both open-source and private software, and the SBOMs for competing vendor offerings could be used to help select less vulnerable software. The resulting competition would also put pressure on vendors to reduce likely sources of vulnerability which would help facilitate analysis of whether included software components had characteristics that made them particularly likely to have vulnerabilities not yet known.[76]

In 2021, pursuant to Executive Order 14028 on *Improving the Nation's Cybersecurity*, the NTIA published the "Minimum Elements" for an SBOM.[77] This standard does not require that an SBOM be more than one layer deep, so one would have to compile multiple SBOMs to get a single picture of all the software rolled up in a single vendor offering. It would make sense for Vendors to be required to do this compilation, but for now there do not appear to be any regulations requiring the production of SBOMs by software vendors. The Cybersecurity and Infrastructure Security Agency (CISA) has proposed a related vulnerability exchange (VEX) document that would help disclose the extent to which software products contain identified vulnerabilities.

It appears that currently only the Food and Drug Administration (FDA), acting under a requirement of section 524B(b)3 of the Food and Drug Omnibus reform Act of 2022 (FDORA), has required SBOMs be developed and submitted, in this case for medical devices.[78] Certainly worth addressing financial regulators is the question of how and whether compiled SBOMs should be required for all software in use in financial sector enterprises.

Software engineers have noticed that Large Language Models (LLMs) such as ChatGPT4 do surprisingly well at writing computer code.[79] Since many vulnerabilities arise from quite simple coding errors or sloppiness – for example memory overflow from improper indexing, or scripting function calls that allow arbitrary execution when more limited calls would be safe), there is potential for automated AI-based software editing and testing to reduce the presence of vulnerabilities.[80]

Another approach to reducing the number of vulnerabilities in deployed software is to change the software development process to emphasize security by limiting use of the more dangerous function calls, do more security testing earlier, and streamline architecture—a trend now sometimes called DevSecOps – which inserts Security into the DevOps agile/rapid

[76] The DARPA SocialCyber program has sponsored technology developments that analyze Open Source development project data to locate conditions that indicate likely points for insertion of malicious code. Under this program Margin Research has, for example, developed powerful Artificial Intelligence assisted analysis tools (REAGENT) for this purpose. See also the Linux Foundation's "Open Security Census" project. https://www.coreinfrastructure.org/wp-content/uploads/sites/6/2018/04/pub_ida_lf_cii_070915.pdf and https://www.coreinfrastructure.org/programs/census-program-i/

[77] https://www.ntia.doc.gov/files/ntia/publications/sbom_minimum_elements_report.pdf

[78] https://csrc.nist.gov/csrc/media/Presentations/2023/fda-s-medical-device-program-and-sbom/images-media/JWilkerson-ssca-forum-053123.pdf.

[79] https://www.cnn.com/2023/03/16/tech/gpt-4-use-cases/index.html.

[80] This is the premise underlying DARPA's new AI Cyber Challenge (AIxCC) competition https://www.darpa.mil/news-events/2023-08-09.

development cycle.[81] It would also be possible to streamline the heterogeneity of software by centralizing all calls to a system logger, so that there wouldn't be multiple instances and versions of Log4j installed on the same computer as part of different software packages, making the maintenance and patching process more straightforward.

Resilient Operation Under Attack

Most experts believe that enterprises must learn to operate under constant cyber attack and that at least some level of intrusion is inevitable. Here the emphasis shifts from not just efforts to reduce vulnerability but to a more active style of defense in depth and prepared response.

Centralize network security information and automate incident response. The acronyms SIEM (Security Information and Event Management) and SOAR (Security Orchestration, Automation, and Response) represent software-based solutions offered by various cybersecurity vendors to perform overlapping cybersecurity functions – rolling up security-related indications and alerts from lower-level systems, presenting the data to operators, suggesting possible responses from pre-established incident response playbooks, or even automatically performing certain functions such as stopping execution of apparently malicious processes or quarantining particular desktop computers or servers, preventing them from accessing the network.[82]

Perhaps these systems had more of a role in preventing data breaches in the first place, and perhaps there is room for future improvement. IBM also found that having multiple (fragmented) cybersecurity tools and staffing shortages was associated with an increase in cost, though again not to a significant degree compared to the average cost of a data breach..

A big focus of these SIEM and SOAR systems is reducing the workload on often understaffed and poorly trained cybersecurity staffs by automating low-level functions and enabling a comprehensive view of enterprise cybersecurity indications while also focusing attention on detected incursions and anomalies. It is surprising how small cyber incident response staffs are in many enterprises that have thousands of employees –even those that do incident response completely in house.[83]

The sorts of indications of compromise that these systems can detect include logins from inappropriate IP addresses or locales, excessive file copying or unusual resource accesses, reports

---

[81] https://csrc.nist.gov/Projects/devsecops. IBM and the Ponemon Institute report that enterprises that apply DevSecOps incur substantially lower costs from a cyber breach, but it is unclear whether this is a result of DevSecOps or just that organizations that claim to use this framework are already more sophisticated – or perhaps that there is just one or a few data points and so it is a meaningless result. Cost of a Data Breach 2023, previously cited. Also see https://1665891.fs1.hubspotusercontent-na1.net/hubfs/1665891/A%20Framework%20DevSecOps%20Guide%20Making%20It%20Happen.pdf.

[82] IBM's 2023 Cost of a Data Breach survey suggests that these tools, together with employee training and widespread use of encryption all were correlated with more than a $200,000 decrease in the cost of a data breach – but with the average data breach cost in the study pegged at $4.69 million even the largest effect was on the order of 5% of the average cost – not a very convincing effect.

[83] In a recent survey, 73% of enterprises with 1,000-2,499 employees and 65% of enterprises with 2,500-9,999 employees had five or fewer employees assigned to cyber incident response; sixty percent with more than 10,000 employees had 10 or fewer. Cortex by Palo Alto Networks, *State of Security Automation*: 2021, p. 5. https://www.paloaltonetworks.com/content/dam/pan/en_US/assets/pdf/reports/state-security-automation.pdf.

of phishing attacks, indications from the firewall of unusual incoming or exiting traffic, etc. The sorts of actions that can be done automatically include opening cases and prefilling incident report forms, quarantining affected equipment, locking users out of the system if their accounts appear compromised, and initiating pre-planned incident playbook workflows that are appropriate to the situation.

Increasingly these systems use AI pattern recognition and machine learning to recognize suspicious patterns of activities and either present them to operators and take automated steps – at least initial steps – in response. There is a significant improvement in performance to be gained from this detection process being informed by patterns of activity seen in the entire enterprise and across multiple enterprises.

Use Systemic Threat and Vulnerability Intelligence

One of the major cybersecurity successes over the last few years involves organizations like Microsoft's Threat Intelligence Center, Mandiant, and CrowdStrike actively tracking changing attack styles and targets and even more importantly the universe of attack-related information available for purchase and the evolving capabilities and tactics, techniques, and procedures employed by specific "Advanced Persistent Threat" (APT) organizations.

Because these organizations see attacks on many different customers, they have an unparalleled view of the active threat and vulnerability landscape. Here Microsoft has a unique opportunity to monitor emerging threats and vulnerabilities because of its position as the dominant enterprise personal computer operating system supplier.[84]

Even were the 8,000 members claimed for the "Microsoft Threat Intelligence community" an exaggeration of the number of direct Microsoft FTEs focused on threat intelligence and remediation, it clearly outweighs the handful or at most tens of employees that may have incident detection and response responsibilities in a typical enterprise. Moreover, Microsoft's reach extends well beyond the relatively well-protected desktops in large enterprises to those in less-defended small enterprises and even stand-alone personal computers. As a result Microsoft may see trends in exploits and distributed attacks before they would be seen by enterprise-only vendors.

Many examples of how Microsoft uses the information reported from deployed software including Microsoft Defender 365 are available on the Microsoft Threat Intelligence Center Blog. One example, selected without much search effort, reflects many of the themes already developed in this paper. In outline offerings, Microsoft 365 Defender telemetry alerted investigators to new processes being spawned anomalously from the SolarWinds Serv-U file transfer process.

Further investigation found a vulnerability in an SSH encrypted information exchange and remote execution module. Microsoft provided information to SolarWinds that allowed them to develop a patch, but in the meantime Microsoft updated Microsoft Defender to detect the threat

[84] Microsoft claims that their Microsoft Threat Intelligence community is made up of more than 8,000 world-class experts, security researchers, analysts, and threat hunters analyzing 65 trillion signals daily to discover threats and deliver timely and hyper-relevant insight to protect customers. This research covers a broad spectrum of threats, including threat actors and the infrastructure that enables them, as well as the tools and techniques they use in their attacks. https://www.microsoft.com/en-us/security/blog/topic/threat-intelligence/.

and notified customers in a higher service class and worked with them to contain it. The attacker was identified as a Chinese group that focuses on the U.S. Defense Industrial Base and has also been known to attack via VPNs and Router software.[85]

Combining active information on penetrations and exploits with a focus on understanding the identities, aims, capabilities and Tactics, Techniques, and Procedures (TTPs) of attacker organizations enables them to rapidly understand emerging attacks and to work with software suppliers and organizations under threat to contain and the threat and help provide software patches or other solutions.

In this way Microsoft was able to help blunt Russian cyber attacks on Ukraine.[86] Vendor security offerings that take this sort of threat information into account, especially if combined with machine learning to recognize evolved attack signatures, are likely to be more powerful than detection tools that do not take advantage of such information.

Exercise and Test. Enterprises that routinely challenge their cybersecurity defenses and that practice incident response have been shown to be more secure than others, though part of this effect may be due to these companies being better resourced and better managed than ones that don't. In the future Large Language Models (LLMs) may be useful in establishing realistic threat scenarios. Financial sector regulators should consider requiring cyber audits and cyber stress tests using red teams, in an analogy to the market risk stress tests that are required of banks to review the adequacy of their capital reserves.

From Back-up and Restore to Ensure Continuity of Data Access. The ability to restore data from back-ups can help get systems back up and running and help deal with data integrity issues. Clearly it is not possible to fully restore confidentiality once data has been exfiltrated or subject to unauthorized inspection. The recent rise of attacks involving persistent presence in networks has raised the specter that back-up processes could be corrupted and back-up data could be rendered unusable, increasing increased interest in back-ups on permanent media but also end-to-end encrypted back-up in the cloud and "Back-up as a Service" (BaaS). Rather than the notion of back-up as a snapshot, a more robust solution to ensure data integrity may be storage that integrates multiple storage locations and tracks changes in data state in between back-up snapshots.[87]

The Evolving Balance Between Vulnerability and Defense: Focus on Robust Resilience

On balance, it is possible to be at least somewhat optimistic that basic cybersecurity efforts have kept up with some of the more common and lower level threats, despite dramatic changes in the threat universe. This has taken place despite massive stress on the IT infrastructure from the

[85] See the original report at https://www.microsoft.com/en-us/security/blog/2021/07/13/microsoft-discovers-threat-actor-targeting-solarwinds-serv-u-software-with-0-day-exploit/ and a subsequent very clear analysis at https://www.microsoft.com/en-us/security/blog/2021/09/02/a-deep-dive-into-the-solarwinds-serv-u-ssh-vulnerability/. See also Microsoft's *Digital Defense Report 2023*. https://www.microsoft.com/en-us/security/security-insider/microsoft-digital-defense-report-2023.

[86] https://www.nytimes.com/2022/02/28/us/politics/ukraine-russia-microsoft.html See also *2023 Crowdstrike Global Threat Report*.

[87] See the 2023 Gartner Magic Quadrant report on *Enterprise Backup and Recovery Software Solutions*. https://www.gartner.com/doc/reprints?id=1-2ELVPL9Z&ct=230801&st=sb

blossoming of remote work during the COVID pandemic, dramatic increases in the diversity and complexity of software supply chains and the move to the cloud.

In the future, however, there is a pressing need to be concerned that further developments in technology and in threat evolution, particularly the prospects of major hostile cyber attacks that have not been a part of the discussion, do not upset the apparent balance.[88] Two technological developments need to monitored closely for their implications for cyber security, and government agencies such as NSF and DARPA should fund research to prevent a technological surprise in these areas.

For the last 30 years, essentially all cyber security has been based on public key cryptography – whether the security of data transfer, the ability to store passwords securely, or encrypting data at rest. Quantum computing may hold the possibility of increasing the ability to crack codes as to endanger all current cryptography; but it may also hold the promise of even stronger encryption that even quantum computing couldn't crack. Both basic and applied research is needed in this area, and also monitoring research being conducted in other countries on these problems.[89]

Less dramatically, but more immediately, the increasing capabilities of Large Language Models may join other AI disciplines such as machine learning in affecting the balance between cyber offense and cyber defense. LLMs are already attracting interest for finding and fixing vulnerabilities in computer code, and are already increasing the ability of hackers with poor language and people skills to launch improved phishing attacks, and deep fakes threaten to complicate even more the ability to arrive at truth on the Internet.

[88] See Nicholas Rostow and Abraham Wagner, *Digital Pearl Harbor: Responses to the Growing Threat* (New York: Margin Research, September 20, 2023).

[89] See Lily Ablon, et al., *Going Dark: Implications of an Encrypted World* (Los Angeles: Center for Advanced Studies on Terrorism, 2017).

4. Threat Model – Why the Nation-State Threat is Different

Early History of Hacking

At least one writer traces the history of hacking back to the late 1800s when computer systems didn't yet exist, although there were enough similarities to the early days of exploits to make a plausible comparison between telephone line switching and modern cybersecurity.[90] In the current world there are numerous examples of hackers using telephones to create exploits in technical systems. By the mid-1960s, as modern communications and information technology evolved, hacking as it is now known actually began.

By the mid-1900s, the term hacking was used at MIT and referred to train sets, not computers when students began altering train sets and they became known as hackers.[91] The practice later became associated with changing the technical aspects of a computer to alter it from its intended usage. In the 1960s hacks, again associated with telephones, was known as phreaking, where an individual uses a high pitch noise to trick a phone into receiving operational commands and changing the behavior of the telephone system.

With AT&T's early dominance of the telephone systems; the development of the Unix operating system; and, the invention of the ARPAnet there was a new focus on this emerging technology and efforts to deal with the emergence of malware and other cybercrime techniques. Hacking as it is now known began in the early 1970s, after the popularization of early computers.

As government agencies adopted these new technologies, the Air Force undertook an initial test of their systems in 1971.[92] These teams of technical specialists became known as "Tiger Teams" and were one of the earliest types of hackers. Other examples of hacking can be seen such as with the invention of the world's first computer worm and subsequent antivirus.[93]

More agencies started to test their network security in response to some of the earliest forms of computer hacking, and by the 1980s hacking prompted Congress to enact the Computer Fraud and Abuse Act (CFAA), which still remains a topic of debate. Computers became more popular and mainstream with major corporations, and with personal computers becoming

[90] The first hacking incident is reported to have taken place in 1878 shortly after the invention of the telephone when young phone operators pranked callers by switching telephone lines, exploiting a technical loophole in the system. Jacob Fox, "A Brief History of Hacking," *Cobalt* (December 5, 2022).

[91] "The History of Hacking," *Fortra* (August 17, 2016).

[92] Edward Hunt, "US Government Computer Penetration Programs and the Implications for Cyberwar," *IEEE Annals of the History of Computing* (2012).

[93] Vikki Davies, "The history of cybersecurity," *Cyber* (October 4, 2021).

widespread as did hacking. In the late 1990s during the dot com boom, the nation also saw a rise in hacking, as credit card fraud and illegal wire transfers boomed.

Cybersecurity became increasingly significant with increased reliance on computer systems, the Internet and wireless network standards such as Bluetooth and Wi-Fi, as well as the rapid growth of smart devices and the various connected devices in the "Internet of things." Even these small devices can be hijacked to form bot armies that can be ordered to execute Distributed Denial of Service (DDoS) attacks, or perhaps to exfiltrate critical network data.

1970s: ARPANET. Cybersecurity began in the 1970s when researcher Bob Thomas created a program called Creeper that could move across the ARPAnet leaving a trail wherever it went. Ray Tomlinson, the inventor of email, wrote the Reaper which chased and deleted Creeper, and was the very first example of antivirus software.

1980s: Birth of Commercial Antivirus. 1987 marked the birth of commercial antivirus products. Andreas Lüning and Kai Figge released their product for the Atari ST – which also saw the release of Ultimate Virus Killer in 1987. In the same year John McAfee founded McAfee and released VirusScan.

1990s: The World Goes Online. As the Internet became more available, people began putting large amounts of personal information online. Criminals entities saw this as a revenue opportunity and began stealing data from people and governments via the web. By the middle of the 1990s, network security threats had increased exponentially. Firewalls and commercial antivirus software were produced on a mass basis to deal with the threats of the time.

Towards the end of the decade, two series of attacks symbolized the paradoxical nature of the threat environment that to some extent continues to this day. In 1998, in what was called SOLAR SUNRISE, multiple institutions came under attack just as the U.S. was deploying a small number of forces to Iraq, raising alarms at the nascent National Infrastructure Protection Center (NIPC). Ultimately law enforcement identified the attackers – two California teenagers, who had technical advice from a somewhat older Israeli. When asked why, they said they "did it for the power." [94]

Another attack, underway at the same time, was far more serious. The episode is generally remembered under the code word for the investigation, MOONLIGHT MAZE. Exploiting weaknesses in the Unix-based Sun OS4 operating system, a group operating on behalf of the Russian government exfiltrated sensitive data and documents from U.S. universities and Government agencies including the Departments of Defense and Energy, the atomic weapons labs and NASA, that, if printed out and stacked, would be taller than the Washington Monument. [95]

2000s: Threats Diversify and Multiply. In the early 2000s criminals began to heavily fund professional cyberattacks as governments began to stop large scale hacking, with more serious

[94] https://nsarchive.gwu.edu/briefing-book/cyber-vault/2023-02-28/solar-sunrise-after-25-years-are-we-25-years-wiser.

[95] https://carnegieendowment.org/2016/12/13/russia-and-cyber-operations-challenges-and-opportunities-for-next-u.s.-administration-pub-66433 and https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07180251/Penquins_Moonlit_Maze_PDF_eng.pdf.

sentences to hackers. Information security advanced as the Internet grew, but at a pace well behind the hackers, in part because user organizations lagged well behind best practices in securing their systems and logon credentials—something that continues to be true.

In the first half of the decade the most prominent types of attacks were internet worms affecting the Windows operating system and email address books, DDOS attacks, and website defacing.[96] The second half of the decade saw more hacking oriented toward financial gain and also massive Russian DDOS attacks on the countries of Georgia and Estonia, in association with Russia's invasion of Georgia and displeasure with independent behavior in Estonia.[97] The hacktivist collective Anonymous was created and attacked Scientology.

Finally, at the very end of the decade, in an "ultra sophisticated" complex persistent attack dubbed "Operation Aurora," Chinese hackers exploited a zero-day vulnerability in Microsoft's Internet Explorer to attack Google, Adobe, and dozens of other U.S. technology companies, apparently seeking their source code and other trade secrets, using encryption to disguise the attack and exfiltration.[98]

Developments Affecting Cyber Threats Since 2010

Social Media, Disinformation, and Information Theft for Political Purposes. The rise of social media and the use of social media to spread disinformation is one of the biggest changes in the Cyber landscape since the popularization of the world wide web. Russian interference in the 2016 U.S. election, both through disinformation and by releasing Democrats' hacked emails, may not seem directly relevant to financial institutions, but the financial sector ignores these sorts of attacks at its peril, owing to the importance of confidence to the operation of banks and other financial institutions.[99]

Crypto-Currency Ransom and Crypto-Mining Takeovers. One novel technology that has affected cyber attacks is the rise of crypto-currency such as Bitcoin and Monero. Many wrongly believed crypto transactions to be untraceable, which led to a boom in ransomware and other extortions requesting payment in crypto currency. Another way in which crypto currency has affected cybercrime is the increase in system takeovers with the aim of using CPU cycles for crypto mining, with the mined currency becoming the property of the person or group executing the attack.

Release of NSA and CIA Hacking Tools. No doubt the biggest gifts to cyber attackers has been the release of NSA and, to a lesser extent, CIA "Vault 7" hacking tools and vulnerabilities.

---

[96] https://www.nytimes.com/2009/08/27/technology/27compute.html

[97] http://www.networkworld.com/news/2007/051707-estonia-recovers-from-massive-denial-of-service.html, https://www.nytimes.com/2008/08/13/technology/13cyber.html.

[98] https://www.wired.com/2010/01/operation-aurora/ and https://www.cfr.org/cyber-operations/operation-aurora. In the context of this paper, it is interesting that Morgan Stanley was also hit hard by Operation Aurora, though this was not revealed at the time. https://www.theguardian.com/technology/2011/mar/01/morgan-stanley-chinese-hackers.

[99] See *Senate Intel Releases Election Security Findings in First Volume of Bipartisan Russia Report* (senate.gov). see also Abraham Wagner and Nicholas Rostow, *Cybersecurity and Cyberlaw* (Durham: Carolina Academic Press, 2021).

While patches were released to deal with the underlying vulnerabilities, many systems remained unpatched and vulnerable. Moreover, the way the tools were packaged was itself instructive in how to organize for effective attacks.

Many of the most prominent and damaging attacks of the past decade were built on the NSA tools, including the WannaCry exploit that took down the British National Health Service, the NotPetya attack on Ukraine that incidentally took down the Maersk shipping empire, ransomware aimed at U.S. cities, and a campaign to take over computers to mine Monero.[100] Information on the use of Vault 7 tools is harder to come by. WikiLeaks withheld much sensitive material in the release, but the withheld tools could still have become available to well-placed hackers.[101]

Lack of Effective International Cooperation against Cyber Crime and Theft of Commercial Secrets. The U.S. has been effective in charging, convicting, and imprisoning cyber criminals within the reach of its justice system, and this effectiveness has had a substantial deterrent effect. Hackers in Russia, North Korea, China, and other countries outside of the reach of American justice have been free to commit crimes with impunity, even when they are not engaging in hacks if they are willing to avoid traveling to countries with extradition treaties with the U.S.[102]

Some diplomatic efforts have appeared successful on the surface, but these have not very much affected criminal hacker impunity or long impeded state hacking activities:

- A 2009 visit by the FBI to Moscow, which led eventually to the Russian officials involved being charged with treason.[103]
- An announcement in early 2022, a month before the invasion of Ukraine, that Russia had arrested 14 members of the Revil group and seized its assets that had hacked the U.S. Colonial pipeline, shutting of natural gas supplies in the Eastern U.S. – though the REvil group had apparently ceased operation shortly after the hack. Even more strange, the Revil group seems to have reappeared within three months.[104]

---

abcnews.go.com/Technology/researchers-discover-ongoing-cyberattack-nsa-hacking-tools/story?id=47459684https://www.cnet.com/news/privacy/stolen-nsa-hacking-tool-now-victimizing-us-cities-report-says/

[101] https://www.cybereason.com/blog/blog-wikileaks-vault-7-leak-details, https://www.forbes.com/sites/leemathews/2017/03/08/the-wikileaks-vault-7-cia-dump-shouldnt-terrify-you/?sh=31f68d3f6b8a

[102] See Abraham Wagner and Nicholas Rostow, *Cybersecurity and Cyberlaw* (Durham: Carolina Academic Press, 2021).

[103] *Ibid.*

[103] *Ibid.*

- The 2015 agreement announced by President Obama and Chinese President Xi to limit cyber thefts of intellectual property and to work to fight cyber crime—which did seem to have a positive effect for maybe a year at best.[105]

While Russian President Vladimir Putin claims NATO started the war in Ukraine, it seems unlikely there will be any U.S.-Russian cooperation on these issues. Similarly China spends a lot of time disclaiming its role in espionage and other cyber activities and a resumption of the temporary cyber détente that President Obama and President Xi reached in 2015 does not seem to be a current prospect. Hostile and ungoverned areas in Africa may also serve as refuges for cyber criminals.

*Smart Mobile Phones***.** Smart phones are now the linchpin in multi-factor authentication and, really, in identification generally. It's obvious that more financial transactions occur via smart phones than personal computers. Both Android and iPhones have been subject to various exploits that presumably could be used to steal logon credentials. The phone manufacturers have responded to these threats with frequent updates and by requiring web pages to constantly reload information afresh, to limit cross-site and cross-app attacks. iOS's messaging facility has been notoriously open to hacking.

Apple now offers a lock-down mode for those worried about being hacked, and users are often advised to reboot their phones frequently to allow a fresh copy of the operating system to be installed. Probably it's a mistake to use the phone for sensitive authentication purposes, because it has to be relatively open to do what consumers want it to. It may be wise to use one phone for financial and other sensitive operations and another for surfing the web and watching tik toks and the like.

Proliferation of Internet of Things (IoT) Devices. The number of IoT devices has doubled over the last five years, and increased by more than 20 times since 2010.[106] The typical IoT device has dramatically more processing power and memory than was typical five years ago. This means that the DDOS potential of these devices has dramatically increased, if they are left unsecured as is often the case.

Rise of the Cloud. The rise of the cloud presents opportunities for more professional security management but also additional vulnerabilities, including weakness in one customer's access control leading to access to other customers' data as well. This appears to have happened in the massive and persistent "Operation Cloudhopper" attacks and data exfiltrations by the Chinese APT10 group on Hewlett Packard Enterprise Company and other cloud providers between 2015 and 2019.[107]

Crowdstrike's annual threat report tracks changes in attack methodologies, targets, and purposes. Throughout 2022, cloud-conscious actors primarily obtained initial access to the cloud

[104] U.S.-China Joint Presidential Statement on Climate Change | whitehouse.gov (archives.gov).

[106] https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/ and https://blog.pcloud.com/wp-content/uploads/2023/02/Statista-Number-of-IoT-and-non-IoT-devices-comparison-2010-2025.png.

[107] https://insights.sei.cmu.edu/blog/operation-cloud-hopper-case-study/.

by using existing, valid accounts, resetting passwords or placing webshells or reverse shells for persistence after exploiting public-facing applications such as web servers. Once on a machine, actors attempted to gain access primarily through credentials found in files, but also via the cloud provider's instance metadata services (IMDSs).

Since workloads in the cloud are very dynamic and potentially short-lived, most actors established persistence with valid cloud accounts they already possessed or for which they were able to reset the password. Alternatively, if the actor obtained initial access via a web server, they placed webshells or reverse shells on the compromised machine for persistence.

Access: Actors escalated their privileges by gaining access to accounts with higher privileges, either by finding credentials for these accounts or resetting credentials that already existed.

Privilege: To collect data, actors turned to local systems as well as internal information repositories such as code repositories, SharePoint, internal tooling and databases.

Data: To move laterally inside a cloud environment, actors used protocols such as RDP, SSH and SMB; actors with console access also leveraged services such as EC2 instance connect and the Systems Manager Session Manager to achieve this goal.

Lateral: Actors tried to evade defenses by deactivating security products running inside virtual machines. Other actors attempted to masquerade by choosing proxy exits close to expected victim locations or naming newly created virtual machines according to victims' naming scheme.

Defense: Despite industry reports claiming resource hijacking was the most common impact used in 2022, the most ubiquitous impact technique was actually destructive, with actors removing access to accounts, terminating services, destroying data and deleting resources.[108]

Gartner forecasts that almost half of IT spending on system infrastructure, infrastructure software, application software and business process outsourcing will be allocated to cloud service providers.[109]

Dark Web Markets: Exploits and Credentials for Sale. Exploits and hacker tools for sale on the dark web, hacker tools development effort brokered and the sale of login credentials and session credentials including remote desktop protocol credentials that bypass multifactor logon authentication have all made it easier to penetrate corporate networks, including those of financial institutions.[110] As with other markets, specialization makes everyone more efficient and market availability enables attackers lacking specialized skills to hire experienced hackers or their fruits. Specialized "initial access brokers" steal credentials using advanced tools and then sell either the credentials or the access itself as a service to attackers.[111]

---

[108] *op. cit*., p. 15.

[109] https://www.gartner.com/smarterwithgartner/cloud-shift-impacts-all-it-markets

[110] For a personal interest story about brokered hacker tool development, see www.wired.com/story/confessions-marcus-hutchins-hacker-who-saved-the-internet/ .

[111] The price of access seems to have been declining on the dark web, to as low as $1,500 for high quality credentials, indicating either that access is not as much a constraint as before, or, more likely, that the lucrative

Hacking as a Service, including Ransomware, Phishing, and Malware. Even unsophisticated attackers can avail themselves of sophisticated tools and even zero-day exploits by essentially renting them online – through interfaces that retains the code in the machines or cloud *used by* the entity. This vastly increases the number of capable threat actors and thus the number of enterprises likely to be attacked. Phishing as a Service can pre-fill email addresses and appropriate corporate logos to help make convincing phishing emails on a mass scale.[112] Ransomware as a Service customers typically may combine credentials purchased from access brokers with the Ransomware as a service to effectuate sophisticated attacks without much hacking skill at all.[113]

Inclusion of Cyber in Military Doctrine: Hybrid Warfare. Both Russia and China have integrated information operations and cyber attacks into their military doctrine and have military elements devoted to preparing the cyber battlefield and executing cyber operations leading up to and as part any military campaign. Although perhaps a logical extension of previous experience, this has created new capabilities, escalation ladders, and options for coercive diplomacy. As with any novel situation, it may also engender miscalculation on the part of the aggressor, the defender, or potential intervening parties.

Hobbyists and Hacktivists. In recent years the threat to enterprises from hobbyists and hactivists have receded as enterprises have become more sophisticated. Penalties for cybercrime play a role in reducing these threats from domestic parties. Nevertheless enterprises should continue to be concerned about the hactivist threats owing to intensity of feeling over such issues as climate change and international conflict such as current conflicts (Russia/Ukraine, Hamas/Israel). State actors and their proxies may launch attacks under a hactivist banner. Both the hobbyist and hactivist threats are likely to be exacerbated by credentials and accesses, hacker tools, and offerings on the dark web.

Early in 2023, financial services websites reportedly represented 7% of the total targeted by hactivists.[114] A previously unknown organization, Anonymous Sudan, targeted SAS and other enterprises supposedly in retaliation for the burning of a Quran in Sweden, but the groups also appears to have connections to Russian policy aims and cyber threat organizations. Over the course of the year it also has attacked Microsoft Outlook on mobile devices with a DDOS attack that for a time prevented access by many users, health care organizations in Australia, and the European Investment Bank.[115] Increasingly, groups with both criminal and national policy motivations find it useful to represent themselves as virtuous hactivists.

---

activity of stealing credentials has attracted more market participants and enterprise defenses have not kept up with evolution of infostealer techniques. Advertisements by access brokers more than doubled from 2021 to 2022. See https://go.recordedfuture.com/hubfs/reports/ta-2023-0302.pdf, pp 22-24. See also *2023 CrowdStrike Threat Report*, p.9.

[112] https://thehackernews.com/2023/05/new-phishing-as-service-platform-lets.html.

[113] https://www.microsoft.com/en-us/security/blog/2022/05/09/ransomware-as-a-service-understanding-the-cybercrime-gig-economy-and-how-to-protect-yourself/.

[114] https://www.radware.com/security/threat-advisories-and-attack-reports/hacktivism-unveiled-april-2023/

[115] https://flashpoint.io/blog/anonymous-sudan-ddos-timeline/.

State or criminal actors posing as hactivists should not be ruled out from developing and deploying novel sophisticated attacks, or a brilliant teen-age hacker figuring out a new exploit. Typical hobbyists and hactivists do, however, tend to stay close to the available state of the hacking art, perhaps with some modifications.

Criminal Cyber Operations

Criminal groups are interested in rapid return on investment. While they may innovate, their innovations are likely to be incremental or combinations of available tactics, techniques, and procedures. They are unlikely to sit on zero-day vulnerabilities they come across or sit in a network for months or years before taking an action that produces a financial gain. If they can continually or frequently exfiltrate data such as credentials or credit card numbers that makes them money, or if they are using cpu cycles to mine crypto-currency, they may maintain a persistent presence.

This is quite different from a nation state actor either conducting espionage or creating conditions for an eventual disruptive cyber attack during an international crisis. Such actors may well either maintain persistent presence or just verify a zero-day exploitation works, return occasionally to check on it, but otherwise do nothing.

The main trends affecting the criminal threat are

- Increased support for criminal hacking from a support ecosystem that provides logon or session credentials for initial access, and packaged or online hacking tools including ransomware;
- The degree to which payment methods are either untraceable or allow balances to disappear beyond national borders where criminals are protected; and
- Permissive sanctuary or even support, technical and otherwise, from hostile nations, and, in the case of North Korea, criminal activity as a way of funding the state budget.

Typical ways criminal cyber operations attempt to make money include:

- Extortion via denial of service attacks;
- Ransomware – classically involving disabling systems and encrypting data until a ransom is paid;
- Exfiltration of commercial data, including sensitive customer data, and threatening to release it if a ransom is not paid;
- Theft of data or information that can be sold or used for blackmail, including credit card numbers and security codes, commercial secrets, logon credentials, personal identifying information, embarrassing personal information and pictures;
- Direct theft of financial assets, for example by ordering external transfers from banks, or cryptocurrency balances from exchanges; and
- Theft of computer cycles for crypto-currency mining or to be used for a fee in someone else's bot army.

The most effective criminal attack is the "double extortion" ransomware attack that demands payment both to restore encrypted data and to prevent release of exfiltrated data.[116] Another possibility would be a privateering model in which hostile nations pay bounties for cyber attacks by criminals that take critical U.S. resources offline or that otherwise inconvenience the American people and make them feel that the government doesn't or can't protect them.

Iranian Cyber Operations

According to the Director of National Intelligence:

> Iran's growing expertise and willingness to conduct aggressive cyber operations make it a major threat to the security of U.S. and allied networks and data. Iran's opportunistic approach to cyber attacks makes critical infrastructure owners in the United States susceptible to being targeted by Tehran, particularly when Tehran believes that it must demonstrate it can push back against the United States in other domains. Recent attacks against Israeli targets show that Iran is more willing than before to target countries with stronger capabilities.[117]

With the United States vowing to "have Israel's back"[118] in response to the current conflict in which Iran-backed Hamas has massacred civilians and taken hostages, Americans need to look hard at Iranian Cyber capabilities that might be deployed, for example in case the U.S. seems to be about to use military force against Iran and its proxies. Iran's cyber operations against the U.S. date at least back to its 2011-2013 DDoS campaign that took down the websites of U.S. Banks.

Operation Ababil locked hundreds of thousands of banking customers out of accounts for long periods of time and resulted in tens of millions of dollars in costs to remediate. An NSA briefing document also made clear the motivation for Operation Ababil: "[Signals intelligence] indicates that these attacks are in retaliation to Western activities against Iran's nuclear sector and that senior officials in the Iranian government are aware of these attacks.[119]

Iranian hackers have since exfiltrated data from U.S. universities and targeted U.S. industrial control systems with the prospect of causing physical damage. Iranian APT groups have engaged in disinformation campaigns and supply chain attacks, and have destroyed data through wiper attacks, sometimes disguised as ransomware. Iran has attacked water systems in Israel, with enough sophistication that Israel felt the need to establish deterrence by responding with a cyber attack that disrupted operations at an Iranian port.

The most destructive cyber attacks assumed to come from Iran have targeted Saudi Arabia. In 2012 the Shamoon malware wiped hard drives of tens of thousands of Saudi Aramco windows-based computers. From November 2016 to January 2017, an updated version destroyed databases and files belonging to Saudi government and businesses, including the Saudi Central Bank.

---

[116] *2023 CrowdStrike Threat Report*, p. 12.

[117] https://www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf, p.19.

[118] Biden statement, October 10, 2023.

[119] https://carnegieendowment.org/files/Iran_Cyber_Final_Full_v2.pdf , p. 30.

In November 2021 the U.S. government warned that Iranian government sponsored APT organizations were targeting the U.S. transportation and healthcare sectors. Already in early 2021, Russia and Iran agreed to cooperate on cyber technology and diplomacy. With Russia's reliance on Iranian drones in its invasion of Ukraine, there is every reason to believe that this cooperation has grown closer, despite tension in the cyber area owing to Russian attackers commandeering Iranian infrastructure to launch attacks. Tehran also has good and improving relations with China, in part based on China's access to Iranian oil, though there has been no evidence yet of direct cooperation on cyber matters. Iran has also been known to share hacker tools with Hezbollah.[120]

North Korean Cyber Operations

According to the U.S. Director of National Intelligence,

> North Korea's cyber program poses a sophisticated and agile espionage, cybercrime, and attack threat. Pyongyang's cyber forces have matured and are fully capable of achieving a range of strategic objectives against diverse targets, including a wider target set in the United States.[121]

Pyongyang probably possesses the expertise to cause temporary, limited disruptions of some critical infrastructure networks and disrupt business networks in the United States. North Korea's cyber program continues to adapt to global trends in cybercrime by conducting cryptocurrency heists, diversifying its range of financially motivated cyber operations, and continuing to leverage advanced social engineering techniques.

Beyond Pyongyang's cybercrime efforts, cyber actors linked to North Korea have conducted espionage efforts against a range of organizations, including media, academia, defense companies, and governments in multiple countries. North Korea continues to conduct cyber espionage to obtain technical information almost certainly intended to advance Pyongyang's military and WMD programs.

The plurality of North Korean attacks remain focused on espionage, primarily in South Korea and nearby Asia. DDoS and destructive attacks for political purposes abroad and complex attacks to fund the state budget are other prominent focus of North Korean cyber attacks. The most widespread and destructive North Korean attack, the 2017 *WannaCry* ransomware worm, took advantage of the NSA Eternal Blue exploit released by the Shadow Brokers, locking up hundreds of thousands of unpatched Windows systems, and would have been much more destructive had the code not included a mysterious "kill switch" discovered by hackers.

---

[120] In one heist in 2022, Pyongyang stole a record $625 million from a Singapore-based blockchain technology firm This account in the preceding paragraphs is drawn from https://www.fdd.org/analysis/2022/10/28/the-dangers-of-irans-cyber-ambitions/, https://carnegieendowment.org/files/Iran_Cyber_Final_Full_v2.pdf, https://www.sipa.columbia.edu/how-does-iran-conceive-cyber-part-its-national-strategy, https://iranprimer.usip.org/blog/2023/may/03/report-iran-accelerates-cyberattacks, https://www.mei.edu/publications/iranian-apts-overview, https://besacenter.org/iran-cyber-threat/, https://www.ncsc.gov.uk/news/uk-and-allies-expose-iranian-state-agency-for-exploiting-cyber-vulnerabilities-for-ransom-operations, and https://www.cybercom.mil/Media/News/Article/2945592/iranian-government-sponsored-actors-conduct-cyber-operations-against-global-gov/ .

[121] https://www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf, p.21.

This and other details of the code raised suspicion that destruction and perhaps embarrassing the NSA was the point, rather than actual economic gain.[122] Other North Korean attacks have been much more targeted and sophisticated.

Espionage has been focused on U.S., South Korean, and Japanese military plans and tecnological secrets affecting the military balance on the Korean peninsula. North Korean hackers stole the South Korean/U.S. war plan for a war on the Korean Peninsula, which was carefully guarded. In an attack in September 2016, North Korean hackers infected 3,200 computers, including 700 connected to the South Korean military's internal network, which is normally "air-gapped" or cut off from the internet. The attack even affected a computer used by the defense minister.

Destructive attacks on financial institutions included an attack on Nonghyup Bank in South Korea, beginning in 2010, that culminated in destroying nearly half of the banks servers and paralyzing the bank's computer network for a week in 2011.[123] Another major focus has been criminal activity – harvesting cash from ATMs, stealing bitcoin from crypto exchanges, and even inserting false transfers from the New York Federal Reserve Bank on the part of the Central Bank of Bangladesh.

These attacks are primarily conducted by more than 6,000 hackers, mostly resident in groups not in North Korea but in China, Southeast Asia, and even Europe. The most destructive focused attack by North Korea was the hack of Sony Pictures, conducted out of pique about a comedy film about an assassination of Kim-Jong-Un.[124]

Details of these attacks show a great degree of patience, skill and cunning, both in terms of technical hacking capabilities and in some cases inventing persona or impersonating others and employing ruses in order to trap well-placed bank employees into unwittingly placing malware on their networks. The army of hackers is recruited as if for an Olympic sports team by selecting those with demonstrated math talent at an early age.[125]

---

[122] https://www.wired.com/2017/05/wannacry-ransomware-hackers-made-real-amateur-mistakes/ and https://www.wired.com/story/confessions-marcus-hutchins-hacker-who-saved-the-internet/. For attribution of WannaCry 2.0 see https://www.justice.gov/opa/pr/north-korean-regime-backed-programmer-charged-conspiracy-conduct-multiple-cyber-attacks-and.

[123] https://www.koreatimes.co.kr/www/news/nation/2011/05/117_86369.html.

[124] Department of Justice, *North Korean Regime-Backed Programmer Charged With Conspiracy to Conduct Multiple Cyber Attacks and Intrusions* (September 8, 2018).

[125] See The Incredible Rise of North Korea's Hacking Army | *The New Yorker*, The Evolution of North Korean Cyber Threats The Asan Institute for Policy Studies | The Asan Institute for Policy Studies, Understanding the Past, Present, and Future of North Korean Cyber Operations | Belfer Center for Science and International Affairs, *Why Is North Korea So Good at Cybercrime?* – The Diplomat, Assessed Cyber Structure and Alignments of North Korea in 2023 | Mandiant, Mapping North Korean Cyber Strategies - RSIS, CO22136.pdf RSIS NK cyber 2022, China, North Korea pursue new targets while honing cyber capabilities - Microsoft On the Issues, https://go.recordedfuture.com/hubfs/reports/cta-nk-2023-0622.pdf, North Korea-linked supply chain attack comes after years of steady, cyber aggression in regime, and North Korea Cyber Threat Overview and Advisories | CISA, North Korea's Cyber Capabilities and Strategy | DGAP, and https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1aFyW, https://www.nytimes.com/2017/10/10/world/asia/north-korea-hack-war-plans.html.

Russian Cyber Operations

According to the Director of National Intelligence,

> The Ukraine war was the key factor in Russia's cyber operations prioritization in 2022. Although its cyber activity surrounding the war fell short of the pace and impact we had expected, Russia will remain a top cyber threat as it refines and employs its espionage, influence, and attack capabilities. Russia views cyber disruptions as a foreign policy lever to shape other countries' decisions.[126]

Russia is particularly focused on improving its ability to target critical infrastructure, including underwater cables and industrial control systems, in the United States as well as in allied and partner countries, because compromising such infrastructure improves and demonstrates its ability to damage infrastructure during a crisis. The DNI report goes on to separately assess Russia's malign influence operations:

> Russia presents one of the most serious foreign influence threats to the United States, because it uses its intelligence services, proxies, and wide-ranging influence tools to try to divide Western alliances and increase its sway around the world, while attempting to undermine U.S. global standing, sow discord inside the United States, and influence U.S. voters and decision making.

Moscow probably will build on these approaches to try to undermine the United States as opportunities arise. Russia and its influence actors are adept at capitalizing on current events in the United States to push Moscow-friendly positions to Western audiences. Russian officials, including Putin himself, and influence actors routinely inject themselves into contentious U.S. issues, even if that causes the Kremlin to take a public stand on U.S. domestic political matters.[127]

Moscow views U.S. elections as opportunities for malign influence as part of its larger foreign policy strategy. Moscow has conducted influence operations against U.S. elections for decades, including as recently as the U.S. midterm elections in 2022. It will try to strengthen ties to U.S. persons in the media and politics in hopes of developing vectors for future influence operations.

Russia's influence actors have adapted their efforts to increasingly hide their hand, laundering their preferred messaging through a vast ecosystem of Russian proxy websites, individuals, and organizations that appear to be independent news sources. Moscow seeds original stories or amplifies preexisting popular or divisive discourse using a network of state media, proxy, and social media influence actors and then intensifies that content to further penetrate the Western information environment. These activities can include disseminating false content and amplifying information perceived as beneficial to Russian influence efforts or conspiracy theories.[128]

---

[126] www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf.

[127] See Dave Aitel, Sophia d'Antoine, Thomas Garwin, Ian Roos, Nicholas Rostow, Justin Sherman, and Abraham Wagner, *Russia's Cyber Operations: A Threat to American National Security* (New York: Margin Research, 2023).

[128] *Op. Cit.* p. 15.

Russian use of information and the Internet has strong historical roots. For more than a century, Russia has used forgeries, disinformation, and falsehood-propagation as an important military and intelligence tool, using "active measures" (*aktivnye meropriyatiya)* as covert and deniable political influence and subversion operations, from corruption and disinformation to assassination and sponsorship of coups.

This history stretches to the time of the Tsars. Throughout the Communist period, Russia used "active measures" including front organizations and spreading false information. In the context of the Internet and advanced information technologies, the Russians emphasize deniability, blur the lines between public diplomacy and propaganda, and use disinformation as a form of political warfare.[129]

Russia under Putin claims to be engaged in an ongoing, existential struggle against Western and NATO forces threatening the regime. This perception is increasingly driven by paranoia and conspiratorialism, particularly about "color revolutions" that the Putin regime fears within Russia. The Putin regime also uses illegal and asymmetric tactics, such as assassinations and disinformation campaigns, to achieve its aims at home and abroad. Hence, the Russian government sees the Internet and the free flow of information it engenders as both a serious threat and equally serious opportunity.

Russian military theorists avoid using the terms cyber or cyberwarfare, preferring to see cyber operations in the broader framework of information warfare. This holistic concept includes computer network operations, electronic warfare, psychological operations, and information operations. Consistent with Soviet notions of combating ongoing threats from abroad and within, Russia views the "information confrontation" – the struggle over "information space" -- as constant and unending, and largely unaffected by western notions of a division between peace and war or concepts of international law.

Information operations to maintain political control within Russia are continuous with operations abroad to manipulate public opinion, political results, and government actions to be consistent with Russian objectives. Unlike the U.S., which forbids its intelligence agencies from spying on U.S. persons, the Federal Security Service (FSB – successor to the Soviet KGB) is responsible for wiretapping and monitoring Internet traffic within Russia for political purposes as well as engaging in espionage and active measures abroad.

Moscow also views cyber operations as a means of disruption for disruption's sake. It can degrade an enemy's military communications, disrupt a foreign company's operations in Russia, and achieve other objectives by means that do not amount to an overt use of military force and still retain the ability claim plausible deniability.

Offensive cyber operations therefore play a large and increasing role in Russian military operations and strategic deterrence. While the Russian military and intelligence services were slow to embrace cyber operations, the government has made significant investments in the last decade and continues to bolster offensive and defensive cyber capabilities. Russian patriotic hackers, front groups, and cyber-criminal syndicates, added to military and intelligence

[129] See Aitel, et al., *op. cit*.

capabilities, have become central to Russian offensive cyber operations. They provide easily mobilized, anonymous, and deniable cyber assets and actors.

The most infamous Russian information warfare proxy group is the Internet Research Agency (IRA), a troll farm initially based out of St. Petersburg, Russia and funded by the late Yevgeny Prigozhin, head of the private military company Wagner Group. The IRA worked to advance the Kremlin's objectives abroad by, among other things, creating fake news articles and posts and then falsely amplifying them on U.S. social media platforms. It has in the years since opened covert outposts outside of Russia, such as in Mexico and Nigeria, to spread disinformation in the West.

Other commercial operatives, such as Positive Technologies, and Kaspersky, are known to support government requirements, in some cases provide direct support to Russian government operations and engage in the international sale of commercial security products. For example, Positive Technologies annually hosts the largest hacking conference in Russia, which the Russian security services use as a venue to recruit hackers to work for the Russian intelligence community.

While Russia has used hackers and criminal networks in the past, evidence now suggests that they are being augmented, if not entirely replaced, by FSB, SVR (Russian Foreign Intelligence Service, successor to the KGB's First Chief Directorate) and GRU (Russian Military Intelligence Organization), and more tightly associated hacker units. The Russian model includes elements of the intelligence services and "external" contract activity, principally the IRA. Russian military doctrine views cyber operations as a mechanism to disrupt enemy force generation and will to fight and to extend the battlefield to the full depth of the enemy's territory.

Specialized units within and associated with the FSB and GRU have been identified by the United States Government, NATO, and commercial cyber security organizations as being responsible for specific cyber operations and types of attacks, including:

- The "Turla" group, operating for the FSB's Center 16 group in Ryazan has over 20 years of experience, dating back to the Moonlight Maze attacks, and later attacks on air-gapped U.S. military systems, hijacking of satellite communications to exfiltrate data back to the FSB, and hijacking Iranian and criminal botnets for espionage.[130]
- APT29, also known as "Cozy Bear" or "The Dukes," responsible for the SolarWinds attack, operating since 2013, is generally believed to operate on behalf of the SVR, and possibly with the FSB. It has attacked the Democratic and Republican National Committees and the Pentagon as well as other targets in the U.S. and other Western countries. It has used advanced phishing and waterhole attacks as well as IT supply chain attacks using tens of other firms as vectors in addition to SolarWinds, and has recently used Microsoft Teams as an initial attack vector.[131]

[130] https://www.wired.com/story/turla-history-russia-fsb-hackers/, https://attack.mitre.org/groups/G0010/.

[131] As part of the Solar Winds attack it was also inside the Danish Central Bank network for over seven months https://socradar.io/apt-profile-cozy-bear-apt29/, https://attack.mitre.org/groups/G0016/.

- "Sandworm Team" a destructive operation associated with the GRU's Unit 74455, has been active since 2009 or before, attacking Ukraine's electric grid and government agencies in 2015-6, the NotPetya attack in 2017, the 2017 French Presidential Election, 2018-9 attacks on Georgia. The 2018 Olympic Destroyer attacks on the Winter Olympic games in South Korea involved false flag techniques and nearly incapacitated the IT infrastructure for the event.[132]
- APT28 or "Fancy Bear," or GRU Unit 26165, operating since 2004, is notorious for hacking the Hillary Clinton campaign but also has attacked a U.S. nuclear facility, and chemical companies and chemical weapons organizations, the German Parliament, and Ukraine.[133]

Russian cyber operations in the context of the invasion and occupation of Ukrainian territory since 2014 illustrate these conclusions. Their successes and failures shed further light on the status of the balance between cyber offense and defense.[134] Nevertheless, the peculiar circumstances of the 2022 Russian invasion attempting to decapitate Ukraine's government and take over all of the country may have led observers to be overly optimistic because of the seeming limited success of Russian cyber operations in determining the outcome of the operation.

Russian cyber operations were affected by poor preparation, planning and coordination in the same way that conventional forces were: they were not told in advance that they were going to participate in an imminent attack on Ukraine and the whole effort that was hobbled by poor planning and coordination. Moreover, Ukraine was already hardened against cyber attack by eight years of war with Russia in the Donbas and American Government's clear warning of impending Russian attack on Ukraine led American tech firms, prominently including Microsoft, to give substantial technical assistance to Ukraine's government and Ukrainian companies to blunt Russian cyber attacks.

Active monitoring of Ukrainian networks for attack signatures of Russian attacks and rapid technical response to these identified attacks was a substantial innovation that reduced damage. Nevertheless in the early hours of the attack, Russia was able to take down access to the Viasat internet satellite system, radically reducing internet service in Ukraine. In the absence of the

---

[132]Some of these attacks were conducted in conjunction with APT28. https://attack.mitre.org/groups/G0034/, https://www.wired.com/story/untold-story-2018-olympics-destroyer-cyberattack/, https://www.washingtonpost.com/outlook/2019/12/04/we-need-hold-kremlin-responsible-its-cyberattack-olympics/, Andy Greenberg, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (New York: Doubleday, 2019).

[133] https://attack.mitre.org/groups/G0007/.

[134] See Cyber Operations during the Russo-Ukrainian War, Nato-Cyber-Report_11-06-2021-4f4ce.pdf, Russian Cyberwarfare: Unpacking Kremlin Capabilities - CEPA, IF11718 CRS Russian Cyber Units, CFR Tracking Cyber Operations and Actors in the Russia-Ukraine War | Council on Foreign Relations, CEIP- Cyber Operations in Ukraine: Russia's Unmet Expectations - Carnegie Endowment for International Peace, CEIP-What the Russian Invasion Reveals About the Future of Cyber Warfare - Carnegie Endowment for International Peace, Russia behind cyber attack with Europe-wide impact an hour... - NCSC.GOV.UK, What we've learned from a year of Russian cyberattacks in Ukraine - The *Washington Post*, "The role of cyber weapons in Russia's war on Ukraine," : NPR, and Bearing Witness: Uncovering the Logic Behind Russian Military Cyber Operations.

provision of Space-X Starlink terminals this would have had a much greater effect on the course of the war.

The overall weight of Russian cyber and information operations have been focused on subversion and political objectives more than tactical and strategic military goals; the poor performance of cyber offense in the Ukraine context may prompt changes to make it more effective in the context of hybrid warfare against a technically sophisticated adversary. A future surprise attack against a country not as well prepared and well warned as Ukraine could have very different results.

Chinese Cyber Operations

According to the Director of National Intelligence,

> China probably currently represents the broadest, most active, and persistent cyber espionage threat to U.S. Government and private-sector networks. China's cyber pursuits and its industry's export of related technologies increase the threats of aggressive cyber operations against the U.S. homeland, suppression of the free flow of information in cyberspace—such as U.S. web content—that Beijing views as threatening to the CCP's hold on power, and the expansion of technology-driven authoritarianism globally.[135]

If China feared that a major conflict with the U.S. were imminent, it almost certainly would consider undertaking aggressive cyber operations against U.S. homeland critical infrastructure and military assets worldwide. Such a strike would be designed to deter U.S. military action by impeding U.S. decision-making, inducing societal panic, and interfering with the deployment of U.S. forces. China almost certainly is capable of launching cyber attacks that could disrupt critical infrastructure services within the U.S., including against oil and gas pipelines, and rail systems.

China leads the world in applying surveillance and censorship to monitor its population and repress dissent. Beijing conducts cyber intrusions that are targeted to affect U.S. and non-U.S. citizens beyond its borders— including journalists, dissidents, and individuals it views as threats— to counter views it considers critical of CCP narratives, policies, and actions. China's cyber espionage operations have included compromising telecommunications firms, providers of managed services and broadly used software, and other targets potentially rich in follow-on opportunities for intelligence collection, attack, or influence operations.

The DNI report goes on to separately assess China's malign influence operations, stating that "Beijing will continue expanding its global intelligence and covert influence posture to better support the CCP's political, economic, and security goals." China is attempting to sow doubts about U.S. leadership, undermine democracy, and extend Beijing's influence, particularly in East Asia and the western Pacific, which Beijing views as its sphere of influence. Beijing largely concentrates its U.S.-focused influence efforts on shaping U.S. policy and the U.S. public's perception of China in a positive direction but has shown a willingness to meddle in select election races that involved perceived anti-China politicians.

[135] www.dni.gov/files/ODNI/documents/assessments/ATA-2023-Unclassified-Report.pdf

China is now using a sophisticated array of covert, overt, licit, and illicit means to try to soften U.S. criticism, shape U.S. power centers' views of China, and influence policymakers at all levels of government. PRC leaders probably believe that a U.S. bipartisan consensus against China is impeding their efforts to directly influence U.S. national-level policy regarding China. Beijing has adjusted by redoubling its efforts to build influence at the state and local level to shift U.S. policy in China's favor because of Beijing's belief that local officials are more pliable than their federal counterparts.

Chinese actors have become more aggressive with their influence campaigns, probably motivated by their view that anti-China sentiment in the U.S. is threatening their international image, access to markets, and technological expertise. Beijing's growing efforts to actively exploit perceived U.S. societal divisions using its online personas move it closer to Moscow's playbook for influence operations.

China is intensifying efforts to mold U.S. public discourse—particularly by trying to shape U.S. views of sensitive or core sovereignty issues, such as Taiwan, Xinjiang, Tibet, and Hong Kong—and pressure perceived political opponents. As part of efforts to stifle anti-Beijing criticism, the PRC monitors overseas Chinese students for dissident views, mobilizes Chinese student associations to conduct activities on behalf of Beijing, and influences research by U.S. academics and think tank experts.

China is rapidly expanding and improving its artificial intelligence (AI) and big data analytics capabilities, which could expand beyond domestic use.[136] For more than a century, Chinese leaders have sought greater access to technology and information to support their national objectives and military capabilities.

The Chinese Communist Party (CCP) has always understood the importance of controlling information to guarantee its domestic position and maximize its ability to manage competition and conflict. Starting in the 1970s, China moved to acquire technologies in order to collect, store, process, and manage information. Today, the success of the country's cyber and communications development is most visible in areas such as 5G (communications) and artificial intelligence (AI).

The People's Republic of China (PRC) has invested substantial sums in technologies related to surveillance, espionage and cyberwarfare.[137] China's cyber operations reflect major

[136] Aitel, et. Al, *China's Cyber Operations, op. cit*., p. 10.

[137] NSA's Cybersecurity Director Rob Joyce has recently been quoted as saying that "The PRC's goal is developing capabilities to disrupt critical infrastructure in the event of future conflict" in Sydney J. Freedberg, Jr., "Chinese 'Volt Typhoon' hack underlines shifting Beijing's Targets, *Breaking Defense* (June 7, 2023). See also, *China's Cyber Power and Military-Civil Fusion* (New York: Margin Research, February 2023). See also, Dean Cheng, *Cyber Dragon: Inside China's Information and Warfare Operations* (Santa Barbara: Praeger, 2017), Michael Pillsbury, *The Hundred Year Marathon: China's Secret Strategy to Replace America as the Global Superpower* (New York: Henry Holt & Co., 2015), Nick Beecroft, *The West Should Not Be Complacent About China's Cyber Capabilities* (Washington: Carnegie Endowment for International Peace, July 6, 2021), Gordon G. Chang, *The Great U.S.-China Tech War* (New York: Encounter Books, 2020) and Anthony H. Cordesman, *China: The Civil-Military Challenge,* (Washington: Center for Strategic and International Studies, January 4, 2022). See Lyu Jinghua, *What Are China's Cyber Capabilities and Intentions?* (Washington: Carnegie Endowment for International Peace, April 1, 2019) and China State Council, New Generation Artificial Intelligence Development Plan (Beijing, July 2017). See also Katharin Tai and Yuan Yi Zhu, "A historical explanation of Chinese cybersovereignty,"

advances made in these areas. China has shaped and reshaped its national cyber ecosystem, which it exploits in new and innovative ways. The mechanism used by the PRC is their integrated Military-Civil Fusion (MCF) strategy, which has greatly expanded China's cyber capabilities in intelligence, espionage, deception, and cyber warfare.[138]

Extensive facial recognition, monitoring of financial transactions, and personal use of connected devices, such as mobile phones and laptops, and social media and other applications provide the means to use the technology base for information and control. The Chinese government tracks individuals and their behavior. Users can access Chinese sites, and versions of U.S. sites, but the government monitors and controls interactions with servers and sites outside China when capable.

The technology also has enabled espionage operations on a scale never before imagined. Operations include theft of intellectual property, extraction of personal data, and penetration of strategic systems—activities going well beyond the traditional intelligence mission of stealing secrets for national security purposes. China collects vast amounts of data by which accesses protected networks and commercial enterprises to make China more competitive in world markets.

As part of their long-term competition with the United States, the Chinese government and CCP view collection and hoarding of information as an investment in the future. It is a strategic aim, not merely a near term tactic. In the area of cyberwarfare, Beijing, however, looks at cyberspace in the broader context of information space.

The ultimate objective is, not "control" of cyberspace, but control of information, a vision that dominates China's cyber operations. Chinese military strategists have also begun to discuss the emergence of what they refer to as "intelligentized warfare," which includes the use of information analysis and AI technologies to target an adversary's "cognition."

China's cyber capabilities and operations have increased exponentially to the point where they pose a highly significant national security threat to the United States and all China's perceived adversaries. The size of the attack surface exponentially increases the risk from such cyber operations and capabilities. China continues to invest huge sums in this technology path. It is clear that the threat will continue to become even greater than it is now. Knowledge of these details of the Chinese cyber strategy and threat need to become a central part of the U.S. national security discourse with respect to cybersecurity.[139]

Over the last decade, the Chinese government has moved in a concerted way to increase its offensive cyber capabilities, both by increasing the scale and sophistication of government hacking efforts and by asserting intimate control over ostensibly civilian technology organizations and processes – for example by requiring all discovered cyber vulnerabilities to be immediately

---

International Relations of the Asia-Pacific (2022) and Nicholas Lyall, "China's Cyber Militias," *The Diplomat* (March 1, 2018).

[138] *Ibid.*

[139] China has come a long way even since its discovered attempted theft of Google's source code in 2009 led Google to withdraw from China in 2010. https://www.technologyreview.com/2018/12/19/138307/how-google-took-on-china-and-lost/

reported to the government and not otherwise disclosed, so they can then be exploited. China has likely surpassed the United States in automated vulnerability discovery. When an engineer at Alibaba reported the Log4Shell vulnerability to Apache rather than first to the Chinese government, he was punished.

Chinese cyber threat organizations have the capacity to reverse-engineer exploits that have been used against Chinese targets and use them against the West, to stealthily exploit vulnerabilities before they are widely known, and then to dramatically expand the scale and scope of attacks once vulnerabilities do become known. Substantial year-over-year increases in zero-day exploits by China have occurred since 2020.

While China still relies on loosely associated hacker groups for deniability of a government role in intellectual property threats, over the past few years cyber offensive capabilities have been realigned and made more focused and strategic under the control of the Ministry of State Security (foreign intelligence) and the Strategic Support Force of the Peoples Liberation Army. In 2023 press reports have revealed both a massive Chinese penetration of Japan's military computer networks[140] and U.S. critical infrastructure on Guam and elsewhere by the "Volt Typhoon" threat group, with the conclusion that this activity was oriented to slowing or stopping a U.S. response to a possible Chinese blockade or invasion of Taiwan.[141]

All U.S. institutions need to assess their vulnerabilities and manage their risks in light of these Chinese practices, as well as Russian cyber operations.[142]

---

[140] https://www.washingtonpost.com/national-security/2023/08/07/china-japan-hack-pentagon/.

[141] https://www.lawfaremedia.org/article/u.s.-and-partners-release-joint-cybersecurity-advisory-on-volt-typhoon.

[142] See The 5×5—China's cyber operations - Atlantic Council, China Cyber Threat Overview and Advisories | CISA, What Are China's Cyber Capabilities and Intentions? - Carnegie Endowment for International Peace, China's Cyberattack Strategy Explained, China's Cyber Capabilities: Warfare, Espionage, and Implications for the United States, CSA_CHINESE_STATE-SPONSORED_CYBER_TTPS.PDF, How China built a one-of-a-kind cyber-espionage behemoth to last | MIT Technology Review, Cyberwarfare by China - Wikipedia, Emerging Cyber Threats: No State Is an Island in Cyberspace, China cyberattacks are a 'defining threat': top U.S. cyber official, IB-323_China's-Developing-Cyber-Warfare-Capabilities.pdf, What it will look like if China launches cyberattacks in the U.S. - POLITICO 1830481, Taiwan's Offensive Cyber Capabilities and Ramifications for a Taiwan-China Conflict | Council on Foreign Relations, *How China Transformed Into a Prime Cyber Threat to the U.S. - The New York Times*, China's National Cybersecurity Center - Center for Security and Emerging Technology, NSA and Partners Identify China State-Sponsored Cyber Actor Using Built-in Network Tools When Targeting U.S. Critical Infrastructure Sectors > National Security Agency/Central Security Service , Press Release View, China Flaunts Its Offensive Cyber Power - *War on the Rocks*, Recent Chinese cyber intrusions signal a strategic shift | *The Strategist*, China's cyber warfare has grown on the back of civilian recruits, and Experts say China's low-level cyberwar is becoming severe threat | China | *The Guardian*.

5. Reducing Financial Sector Cyber Risk and Cost

Problem of Life in The Digital World

The rapid transition of most U.S. institutions to the digital, connected world makes the multi-faceted national infrastructure subject to catastrophic attack and failure. While this digital revolution was taking place, the realm of cyber "attack" and threat also evolved from bored students who engaged in hacking to major criminal organizations and hostile foreign intelligence services and militaries.

This change in the nature and level of threat was not appreciated as worldwide use of the cyber ecosystem exploded. In part, this failure of understanding can be attributed to the speed in which this threat evolved, and in part to the relatively glacial pace in which the federal bureaucracy adapts to changing missions and direction of resources to meet such rapidly evolving threats.[143]

Meeting this challenge and enhancing the security of this infrastructure requires changes in the national approach to organization and decision-making, improved threat detection and analysis, enhanced resilience, hardening of the cyber infrastructure, and preparation for future cyber disasters. It is still the case that the Intelligence Community has failed to focus effectively on the generation of malicious code and those that are engaged in its development so that major cyber attacks can be detected. Current U.S. policy is, unfortunately, far more focused on "incident reporting" with little in the way of programs to deal with the aftermath of catastrophic attacks.[144]

Well into the first decade of the Internet the sophisticated view was that financial services institutions and other businesses would solve the technical security problems in their self-interest. Insurance companies would enforce adoption of best practices on commercial actors and so heavy-handed government action was not necessary. Other sectors and the government itself would take

[143] Looking back at the overall national intelligence budget (NFIB) in the early 1990s it would be possible to conclude that no subject of either intelligence or law enforcement interest would ever use a cell phone or the Internet. Offices and programs to meet these evolving challenges were largely nonexistent, and the few that did exist had operating budgets best described as trivial. A serious history of this programmatic evolution has yet to be written and would likely run into serious classification problems if ever attempted.

[144] See Nicholas Rostow and Abraham Wagner, *Digital Pearl Harbor: Responses to the Growing Threat* (Margin Research, September 2023), and John Ratcliffe and Abraham Wagner, "U.S. Needs New 'Manhattan Project' to Avoid Cyber Catastrophe," *Newsweek* (May 18, 2022).

a free ride on the security innovations and investments driven by commerce.[145] This assumption turned out not to be true for a variety of reasons.

Rethinking the National Approach to Cybersecurity

After World War II, the United States changed its approach to national security, responding to major changes in the threat environment and the technologies involved as well as to the disappearance of alternative great powers that could maintain the peace. The 1947 National Security Act created the Department of Defense and the Air Force, as well as the Central Intelligence Agency to meet an emerging Soviet threat and the potential use of nuclear weapons by the Soviet Union.[146] Despite the fact that cyber threats merit similarly serious institutional responses, the United States has yet to take analogous action.

The technological "surprise" of the Soviet space and missile program of the 1950s led the President to direct a major review of all U.S. agencies and programs and how the nation would respond to this major change in the threat.[147] As a result. major changes were made in the military services while the Intelligence Community created the National Reconnaissance Office (NRO) as one means to collect much-needed data.

CIA created new offices and centers responsive to the emerging threat, supported by a robust external infrastructure. CIA also undertook substantial organizational and staffing changes to deal with the evolving Soviet threat and produce timely, accurate assessments.[148] The National Security Agency (NSA), newly created within the Department of Defense by Presidential Order in 1952, undertook other needed efforts in the SIGINT area.[149]

In addition, Congress authorized major investments in a supporting analytical infrastructure to assist the Defense Department and the Intelligence Community. It included the national laboratories, federally funded research and development centers (FFRDCs), a substantial number of private-sector contractors, and university centers. Further supporting the research and

---

[145] These include the different situations of different sectors, the capacity of businesses to absorb losses as a cost of doing business, the possibility of zero-day exploits, the prevalence of social engineering attacks, and other factors. The nation now knows better.

[146] Pub.L. 80-253, 61 Stat. 495, enacted July 26, 1947. The 1947 Act created the Central Intelligence Group (CIG) to replace the Wartime OSS intelligence service which was eliminated by President Turman's Executive Order in 1946. The 1948 Central Intelligence Act changed the name to the Central Intelligence Agency. See Charles A, Stevenson, "The Story Behind the National Security Act of 1947" *Military Review* (May-June 2008).

[147] See George B. Kistiakowsky, *A Scientist at the White House* (Cambridge: Harvard University Press, 1976).

[148] Changes within the CIA include the Office of Strategic Research (OSR) created in 1967 as well as the Strategic Evaluation Center (SEC) which proved highly successful over the years. See Robert D. Vickers, Jr., "CIA's Office of Strategic Research: A brief History," *Studies in Intelligence* (March 2018), and Central Intelligence Agency, *National Foreign Assessment Center: Organizational Structure and Functions,* (NFAC Plans and Programs Staff, December 1977, Declassified 2002). Organizational efforts within CIA to deal with cybersecurity and related issues provided far less successful. A history of these efforts, offices and programs remains to be written.

[149] See Thomas L. Burns, *The Origins of the National Security Agency* (United States Cryptologic History, National Security Agency, 1990) (Declassified 2007). In 1959 Congress enacted the National Security Agency Act which provides a separate legislative basis for NSA's activities.

development needed were the creation of ARPA, NSF and funding of university research under the National Defense Education Act.[150]

Improved Threat Detection and Analysis

The DARPA SocialCyber, HARDEN and HAMILTON programs, among others, have demonstrated the possibility of reducing vulnerability to major cyber attack through early detection and analysis of hostile code development. This research has integrated extensive collection of code artifacts, patches and other postings along with the development of a set of AI tools that enable identification of malicious code from a large body of data.

Along with the data collection and AI tool development has been an integrated analysis of the organization, institutions and individuals involved in malicious code development utilizing original source materials and graph databases. It is essential that this strategy and tool set be transitioned to the responsible Intelligence Community elements for use on an ongoing basis.

DARPA was the initial home of the Internet and for decades has engaged in research on related network and software technology. It has long supported Intelligence Community partners at NSA and CIA. Current DARPA programs, including those noted above, have taken an approach that is not currently being applied to the problem by the appropriate executive branch organizations. This integrated approach has several key elements:

- Ongoing collection of open-source data such as public email, code artifacts, communications, and postings from hostile nations (such as China and Russia) that may indicate malicious code to be used in activities such as espionage and cyber warfare.
- Use of AI tools applied to the database to identify both malicious code as well as specific individuals who are "contributors" of potentially malicious code.
- Application of graph database tools to demonstrate links between portions of code and code contributors to allow lateral inferences about other code areas likely to be at risk.
- Integrated use of technical experts who are also involved in the development of offensive cyber tools for the U.S. that understand this software.
- Use of regional experts with native fluency in Chinese and Russian as well as expertise in cyber operations, to examine the supporting infrastructure in hostile nations as well as specific messages in the database.
- Adding a centralized layer to private sector network and endpoint and threat organization monitoring and application of advanced AI machine learning and big-date tools to provide real time protective interventions as attacks are underway.

[150] The National Defense Education Act (NDEA) was passed in 1958 in response to Soviet acceleration of the space race with the launch of the satellite *Sputnik*. The law provided federal funding to "insure trained manpower of sufficient quality and quantity to meet the national defense needs of the United States." In addition to fellowships and loans to students, the legislation bolstered education in the areas of science, mathematics, and modern foreign languages.

- Document storage and processing techniques that are hardened against LLM-assisted fakery.

The United States currently lacks such an integrated approach. It also lacks an alternative approach to detect and counter hostile code development. While the responsible agencies may collect some similar data, they still lack the associated AI tools to evaluate the data on an ongoing basis and do not have programs in place to develop them. This means the United State lacks an effective means to prevent or mitigate a digital Pearl Harbor.

Enhancing Resilience and Hardening of the Cyber Infrastructure

For several decades computer scientists associated with DARPA and elsewhere have looked at issues related to modernizing the Internet and enhancing the resilience of the existing infrastructure. Even though the Internet has gone through a period of explosive growth worldwide that was unimagined at the outset it continues to operate with many of the technologies and protocols developed in the 1960s. Even the current IPv4 Internet protocol is a product of the 1980s.[151] Its successor, IPv6, is still being deployed and lacks the type of resilience that most experts believe is essential.

Resolving this aspect of the problem lies in the hands of several worldwide bodies and outside the control of any U.S agency or intelligence service. Whether the worldwide infrastructure will continue with periodic upgrades such as IPv6 which leave many vulnerabilities unresolved, or whether a new and possibly parallel Internet 2.0 will come into being remains an open question.[152] At the same time, the U.S. government could take steps to build, adopt, and mandate for Americans an Internet 2.0. Given U.S. international influence, such a step should not be ruled out, although it currently lacks broad international support.

Apart from the backbone Internet infrastructure itself, software ranging from operating systems to a myriad of applications remain vulnerable to hostile cyber attack. Important findings from the DARPA SocialCyber analysis as well as other recent research has been that popular operating systems, such as Linux, are vulnerable to "contributors" submitting patches and other changes that may in fact introduce malware into the system. An essential part of the intelligence mission here is to monitor these contributions and their contributors with specialized AI tools developed for this task.[153]

Most operating systems and applications are vulnerable. They therefore should be reviewed for specific vulnerabilities, and there need to be plans made for the event they are

[151] Internet Protocol version 4 (IPv4) is the fourth version of the Internet Protocol (IP) and is still one of the core protocols of the Internet. IPv4 was first deployed for production in 1982 and on the ARPAnet in January 1983 and is still used to route most Internet traffic today, even with the ongoing deployment of Internet Protocol version 6 (IPv6) its successor.

[152] An additional question not often discussed is the physical vulnerability of the infrastructure. The operation of the Internet backbone in the U.S. depends on several core routers that are located in unguarded commercial buildings whose locations are well-known. An adversary could easily destroy some or all of these.

[153] See, for example, Ralph Ramsauer, et. al., "The Sound of Silence: Mining Security Vulnerabilities from Secret Integration Channels in Open-Source Projects," *ACM Proceedings* (November 2020).

attacked. The spate of ransomware attacks in recent years at least provides some insight into how failure at some level might impede organizations' operations, such as that of pipeline companies or hospitals, and suggests what can be done to improve resilience here. Mitigation might take the form of separate servers and software or movement to a more protected cloud environment.[154]

Development of Resilient Computer Code and Appliances

It is important to go beyond a mindset that focuses on cyber defense and even a cyber resilience – which suggests rapidly picking up the pieces after an attack – to a mindset of operating while under attack, with enterprises using technology to detect and contain attacks in process and to ensure that continuity of operations – including data confidentiality, integrity, and access, will be assured under attack.

Technological aspects of this philosophy could include superior identification systems so that all users and processes on sensitive networks are inherently identified and not granted access that they aren't allowed. The current Internet architecture militates against this and the increased use of smartphones as identification only adds a hackable device to the process. One possibility would be to twin the Internet – providing an anonymous venue of web surfing and personal use and another for sensitive uses including financial transactions. Cryptocurrency should be avoided for now, given the proven relative insecurity of existing exchanges.

Another would be backup processes and systems that either use non-erasable media or super-hardened software that provides equivalent protection. But even 100% reliable speedy backup does not address the problem of data theft, which is obviously secure in financial institutions.

To support the development of a more resilient cyber infrastructure one useful concept would be to provide far greater resources for software engineering institutes and similar institutions within the U.S. that would further develop much needed standards and data formats for documents and other digital media.[155] At present, there are virtually no common or accepted standards for either data formats or code development. The lack of standards renders the U.S. cyber ecosystem more vulnerable to hostile attack than otherwise would be the case.

Exactly how this critical task can be accomplished should be a matter of major concern for both the Executive branch and Congress. Just as the nation responded to the Soviet threat and new technologies during the Cold War, the United States needs to create a broader scientific and technical infrastructure that is able to produce computer code that is far more resilient in the evolving threat environment. This infrastructure would also take on tasks such as monitoring the

[154] See Lily Ablon, et al., *Going Dark: Implications of an Encrypted World* (Los Angeles: Center for Advanced Studies on Terrorism, 2017) and *Cloud Encryption, Privacy and National Security: Legal and Political Context* (New York: Margin Research: January 2023).

[155] One effort in this direction has been the DARPA Safe Documents (SafeDocs) to develop novel and verified programming methodologies for electronic data formats. This will help protect against input attacks trying to prevent the flow of untrusted data to vulnerable software; and testing software with randomized inputs to find and patch flaws triggered by maliciously created inputs.

effect of LLMs and Quantum computing on cyber security, and otherwise protecting against technological surprise in this area.

There are multiple approaches to meeting this challenge. As in the past, the United States may be best served by using several simultaneously. Certainly, new institutes at academic institutions can play an important role.[156] Another element might be to expand efforts at the existing FFRDCs and national laboratories.[157] At the same time, the supporting commercial contractor base as well as major firms such as Microsoft, Google, and Oracle, just to name a few, need to become an essential part of the process. Responsible government agencies need to be continuously involved, not only as sponsors, but also as overseers to ensure that the process of code development meets the requirements of the evolving threat.

Meeting this challenge needs to remain a core mission for agencies like DARPA and the research components of the Defense Department, the military services, and the Intelligence Community. Such agencies and offices have the internal structure and program management capability to execute this technical mission and need adequate resources to do so. Successful initiatives need to move to places where they can become operational on a sufficiently large scale to benefit the entire cybersecurity effort.

Longer-Term Measures to Prepare for a Cyber Disaster

Although no major or devastating cyber attack has taken place, the threat remains and will continue to grow. Over the longer-term measures need to be taken in order to prepare the country for the growing possibility of a digital Pearl Harbor and respond effectively in the event it happens. Some of the most important measures include:

- Organizational changes which assign the lead to a department, agency and offices having the legal, technical, and management capability to execute an effective cybersecurity program.
- Development of a supporting infrastructure within the government, as well as an adequately funded external base of national laboratories, FFRDCs, contractors, universities, and commercial infrastructure.
- Explore what AI technologies can do to meet emerging threats to critical sectors to ensure that the U.S. maintains an advantage and the ability to provide the needed resilience.
- Require realistic "stress testing" and "war gaming" of potential high-end cyber attacks.

[156] One excellent example is Carnegie-Mellon University's Software Engineering Institute (SEI), which supported DARPA and others in the computer security area for decades.

[157] Over the past several years there has been some discussion about starting yet another Federally Funded Research and Development Center (FFRDC) to deal with the cybersecurity problem, which has faced considerable resistance within the Congress and elsewhere for good reason. Some additional tasking and funding to RAND, IDA, CNA, MITRE and others would serve the same purpose at a lower cost.

- Provide for the development of resilient data formats and standards for code development through new and adequately funded software engineering institutes and others.
- Work with Congress to ensure that both funding and oversight of essential activities are supported.

Actions Specific to the Financial Sector

Compared to many sectors, financial institutions generally have resources to devote to cyber security. The heavily regulated financial sector has some advantages compared to some other business actors in moving toward a safer cyber posture. Regulators can and should require cyber stress testing analogous to stress testing for capital adequacy in the face of uncertain markets. Federal regulators can force mergers of unsafe institutions.

References

Ablon, Lily, et al., *Going Dark: Implications of an Encrypted World* (Los Angeles: Center for Advanced Studies on Terrorism, 2017)

Aitel, Dave, Sophia d'Antoine, Winnona DeSombre, Isabella Garcia-Camargo, Ian Roos, Nicholas Rostow, Jonathan Smith, Alison Strongwater, Abraham Wagner, and JD Work, *China's Cyber Operations: The Rising Threat to American Security* (Margin Research, 2022)

Aitel, Dave, Sophia d'Antoine, Thomas Garwin, Ian Roos, Nicholas Rostow, Justin Sherman, and Abraham Wagner, *Russia's Cyber Operations: A Threat to American National Security* (New York: Margin Research, 2023)

China's Cyber Power and Military-Civil Fusion (New York: Margin Research, February 2023)

Basel Committee on Banking Supervision (BCBS), "Global Systemically Important Banks: Revised Assessment Methodology and The Higher Loss Absorbency Requirement," (July 2018)

Boer, Martin, and Jaime Vazquez, *Cyber Security & Financial Stability: How Cyber-Attacks Could Materially Impact the Global Financial System*, Institute of International Finance, (September 2017) https://www.iif.com/Portals/0/Files/IIF%20Cyber%20Financial%20Stability%20Paper%20Final%2009%2007%202017.pdf?ver=2019-02-19-150125-767

Bookstaber, Richard and Dror Y. Kenett, *Looking deeper, seeing more: a multilayer map of the financial system*. OFR Brief, 16(06). (2016)

Bookstaber, Richard., Paddrik, M., & Tivnan, B. (2017). "An agent-based model for financial vulnerability." *Journal of Economic Interaction and Coordination*, 1-34.

Bradshaw, Tim, Arash Massoudi, and Kara Scannell, "Bogus terror tweet sparks shares blip," *Financial Times*, (April 2013)

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., and Havlin, S., "Catastrophic cascade of failures in interdependent networks." *Nature*, (2010)

Burrows, Oliver and Katie Low, "Mapping the UK financial system, " *Bank of England Quarterly Bulletin* (2015)

Carnegie Endowment for International Peace, "Timeline of Cyber Incidents Involving Financial Institutions," (2022) https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline

Cloud Encryption, Privacy and National Security: Legal and Political Context (New York: Margin Research: January 2023)

Committee on Payments and Market Infrastructures (CPMI) and International Organization of Securities Commissions (IOSCO), *Guidance on Cyber Resilience for Financial Market Infrastructures*, (June 2016)

Committee on Payments and Market Infrastructures (CPMI) and International Organization of Securities Commissions (IOSCO), *Principles for Financial Market Infrastructures*, (April 2012)

Danielsson, Jon, Morgane Fouché, and Robert Macrae, "Cyber Risk as Systemic Risk," VoxEU.org, (August 2016), https://voxeu.org/article/cyber-risk-systemic-risk.

Deutsche Bundesbank, "Financial Stability Review 2017," (November 2017).

European Banking Authority (EBA), "Guidelines On The Criteria To Determine The Conditions of Application of Article 131(3) of Directive 2013/36/EU (CRD) in Relation to The Assessment of Other Systemically Important Institutions (O-SIIs)," December 16, 2014, https://eba.europa.eu/sites/default/documents/files/documents/10180/930752/964fa8c7-6f7c-431a-8c34-82d42d112d91/EBA-GL-2014-10%20(Guidelines%20on%20O-SIIs%20Assessment).pdf.

European Union Agency for Cybersecurity (ENISA), "WannaCry Ransomware Outburst," (May 15, 2017), https://www.enisa.europa.eu/publications/info-notes/wannacry-ransomware-outburst.

Government Accountability Office, *Critical Infrastructure Protection: Treasury Needs to Improve Tracking of Financial Sector Cybersecurity Risk Mitigation Efforts* (September 2020)

Haines, Avril, *DNI Haines Opening Statement on the 2023 Annual Threat Assessment of the U.S. Intelligence Community*, Director of National Intelligence (April 8, 2023)

International Association of Insurance Supervisors (IAIS), "Global Systemically Important Insurers: Updated Assessment Methodology," (June 16, 2016), https://www.iaisweb.org/page/supervisory-material/financial-stability-and-macroprudential-policy-and-surveillance/file/61179/updated-g-sii-assessment-methodology-16-june-2016.

Kopp, Emanuel, Lincoln Kaffenberger, and Christopher Wilson, *Cyber Risk, Market Failures, and Financial Stability*, International Monetary Fund Working Paper no. 17/185, (2017)

Office of Financial Research (OFR), *Cybersecurity and Financial Stability: Risks and Resilience*, OFR Viewpoint 17-01, February 15, 2017, https://www.financialresearch.gov/viewpoint-papers/files/OFRvp_17-01_Cybersecurity.pdf.

Pozsar, Zoltan, Adrian Tobias, Adam Ashcraft, and Hayley Boesky, *Shadow Banking***,**

Federal Reserve Bank of New York Staff Reports (July 2010: revised February 2012)

Ramsauer, Ralph, et. al., "The Sound of Silence: Mining Security Vulnerabilities from Secret Integration Channels in Open-Source Projects," *ACM Proceedings* (November 2020)

Rinaldi, S., Peerenboom, J. P., and Kelly, T. K. (2001). "Identifying, understanding, and analyzing critical infrastructure interdependencies." *Control Systems, IEEE* (2001)

Rostow, Nicholas and Abraham Wagner, *Digital Pearl Harbor: Responses to the Growing Threat* (New York: Margin Research, September 2023)

Ratcliffe, John and Abraham Wagner, "U.S. Needs New 'Manhattan Project' to Avoid Cyber Catastrophe," *Newsweek* (May 18, 2022)

O'Conner, Tom, Haveed Jamali and Fred Guterl, "Will Putin's Hackers Launch a Cyber Pearl Harbor—and a Shooting War? *Newsweek* (June 18, 2021)

Sommer, Peter and Ian Brown, "Reducing Systemic Cybersecurity Risk," Organisation for Economic Cooperation and Development (OECD), (January 2011)

Wagner, Abraham, Thomas Garwin, Nicholas Rostow, Sophia d'Antoine and David Aitel, *DARPA Cybersecurity Planning: Technologies for Keeping the Nation Safe* (Los Angeles: Center for Advanced Studies on Terrorism, 2018).

Wingfield, Nick, "Miscue Calls Attention to Amazon's Dominance in Cloud Computing," *New York Times*, (March 2017)

World Economic Forum (WEF), "Understanding Systemic Cyber Risk," Global Agenda Council on Risk & Resilience, (October 2016)