A REPORT OF THE
TECHNOLOGY AND PUBLIC
POLICY PROGRAM

Conflict and Negotiation in Cyberspace

February 2013

Author
James A. Lewis

50 YEARS | CHARTING OUR FUTURE

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

A REPORT OF THE
TECHNOLOGY AND PUBLIC
POLICY PROGRAM

Conflict and Negotiation in Cyberspace

February 2013

Author
James A. Lewis

50 YEARS | CHARTING OUR FUTURE

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

About CSIS—50th Anniversary Year

For 50 years, the Center for Strategic and International Studies (CSIS) has developed solutions to the world's greatest policy challenges. As we celebrate this milestone, CSIS scholars are developing strategic insights and bipartisan policy solutions to help decisionmakers chart a course toward a better world.

CSIS is a nonprofit organization headquartered in Washington, D.C. The Center's 220 full-time staff and large network of affiliated scholars conduct research and analysis and develop policy initiatives that look into the future and anticipate change.

Founded at the height of the Cold War by David M. Abshire and Admiral Arleigh Burke, CSIS was dedicated to finding ways to sustain American prominence and prosperity as a force for good in the world. Since 1962, CSIS has become one of the world's preeminent international institutions focused on defense and security; regional stability; and transnational challenges ranging from energy and climate to global health and economic integration.

Former U.S. senator Sam Nunn has chaired the CSIS Board of Trustees since 1999. Former deputy secretary of defense John J. Hamre became the Center's president and chief executive officer in April 2000.

CSIS does not take specific policy positions; accordingly, all views expressed herein should be understood to be solely those of the author(s).

This project was funded by the Smith Richardson Foundation.

Center for Strategic and International Studies

1800 K Street, NW, Washington, DC 20006

Tel: (202) 887-0200

Fax: (202) 775-3199

Web: www.csis.org

CONTENTS

1 Cybersecurity and International Security 1

2 The Military Use of Cyberattack 11

3 Constraining Cyber Espionage 42

4 Engagement, Negotiation, and Stability 52

5 Principles and Conclusions 60

About the Author 62

1 CYBERSECURITY AND INTERNATIONAL SECURITY

What counts are the political and military consequences of a violation, . . . since these alone will determine whether or not the violator stands to gain in the end.

—Fred Ikle, "After Detection, What?" 1961

This report looks at the political-military aspects of cybersecurity and attempts to place it in the larger context of international security. Networks are embedded in our economies and our political and social life. They have become the central tool for human activity. These networks form cyberspace. They hold information of immense value, and they control the machinery that provides critical services. They create immense economic benefit, but they are also a major source of risk to nations. Governments have been hesitant to interfere with the golden economic machine, and the result is a weakly governed space, much like a failed state or contested terrain.

Because of the newness of technology, the lack of explicit agreement among states, and rampant cyber espionage and cybercrime, this unstable environment invites miscalculation, misinterpretation, and inadvertent escalation of conflict. Changing this requires identifying which instruments of statecraft are most effective and where we may need new institutions, norms, and laws. Progress in cybersecurity requires manipulating complex international processes to change what governments consider as acceptable national behavior in cyberspace.

Cybersecurity has been an issue for national security since the 1990s, but the U.S. response has been ad hoc and reactive, marked by uncertainty over how to deal with a major new problem for international security. This report identifies six principles that should guide the United States in developing a strategic approach:

- Cyberspace is not a unique environment. States will behave in this environment as they would in any other.
- We cannot "disarm" in cyberspace, and there will be no "global zero" for a cyberattack.
- We have entered a period of sustained, low-level competition for influence where opponents' miscalculations and misperceptions are a source of risk to the United States.
- U.S. interests are best served by embedding cyberattack and cyber espionage in the existing framework of international law, and long-term U.S. interests are best served by winning international agreement to this.
- America's immediate goal in negotiation should be to increase the risks of launching a cyberattack or engaging in malicious cyber activity for both state and nonstate opponents.

- There is a limit to what negotiation can achieve in reducing risk; there will always be risk. The U.S. goal should be to decrease and bound this risk as part of its larger efforts to strengthen international security.

The key to greater security is to establish consequences for hostile action in cyberspace. The greatest weakness in American cybersecurity policy has been the failure to respond to repeated hostile action by foreign powers. This signals that Americans are either indifferent or inept. The issue is to identify the policies and actions that will reduce the incentives for cyberattack and cyber espionage and to build international support for a more stable and secure cyberspace. Hostile actions in cyberspace must have political and military consequences.

Implicit norms for state behavior, derived from international practice, already constrain malicious cyber actions, but these implicit norms are inadequate as a foundation for international security. To date, the calculations of America's foreign opponents, based on their perception of U.S. policy and behavior, are that the benefits of cyber actions outweigh the risk. The United States has inadvertently created a world where cyber espionage and cybercrime are largely risk free. Changing this will require explicit norms for state behavior in cyberspace, reinforced with explicit understandings on outcomes and responses, using direct messages, tacit signals, and observable military capabilities.

The United States cannot be said to have a strategy to deal with cybersecurity, although several documents bearing the title "Strategy" have been issued by various administrations. There are several reasons for this, the most important being that the United States is hampered by a seductive idea that has dominated American politics for the last 30 years—that the role of government should shrink and that the private sector and market are best placed to solve public problems (thus eliminating the need for taxes). This may have been true in the 18th century, but it is not true now in an interconnected, high-technology security environment. Despite having a 2003 National Strategy and a 2009 Cyberspace Policy Review, the United States has not made the necessary strategic calculations.

Concepts from the Cold War Grand strategy continue to echo windily in current debates, but for cybersecurity, the Cold War is a distraction. Some hope to expand strategic deterrence in cyberspace, using U.S. military capabilities to discourage opponents from cyberattack. Although the United States has one of the most powerful cyberattack capabilities in the world, it does not deter other nations from engaging in cyber espionage and cybercrime. Cold War deterrence rested on a framework of norms and understandings (some implicit) between two opponents regarding the use of weapons, "signaling" to indicate concern or rising tensions, and thresholds or redlines to constrain and manage conflict. This framework does not exist for cyberattack.

Changing opponents' calculus of the utility and consequences of cyber exploits and cyberattack cannot be achieved by a deterrent strategy. The immediate dilemma is that America faces potential attackers with differing risk tolerances and uneven abilities to calculate risk. New opponents like Iran and North Korea may be more tolerant of the risk of reprisal or international condemnation, or less able to correctly assess risk. This is also the case for nonstate actors like jihadis and political activists, who are likely to have even higher tolerance for risk. In both cases, risk is asymmetric—they can do far more harm to the United States using cyber techniques than the United States can do to them. We cannot make a credible threat of unacceptable damage.

Cyberattack by Russia or China is unlikely outside some larger conflict. Russia and China will be more responsive to a strategy that relies on the threat of military retaliation to prevent cyberattack, but this clearly does not deter espionage and crime. That said, their use of cyber exploits for espionage and crime is largely unbounded. This is a primary source of risk to the United States because these two nations have the advanced cyber capabilities, are hostile, and are becoming more assertive militarily. Developing techniques that constrain Russian and Chinese cyber activities is a priority for national security.

More than a decade of experience illuminates the requirements for strategy. A "homeland defense" approach is inadequate because securing a global network is a problem for international security. Voluntary private action, where the disaggregated and uncoordinated actions of companies are pitted against powerful and unscrupulous state actors, is also inadequate. The many successful attacks against companies highlight the difficulties for any private-sector actor to successfully defend against foreign military and intelligence services. Homeland security and voluntary action are inadequate for dealing with the threats of a global network because they allow an attacker to defeat any defense in detail—the attacker can concentrate its resources on a single target, overwhelming it in the process.

A purely defensive strategy will not work. Porous technologies, uncertain politics, and human nature combine to give the advantage to the attacker. There is unlikely to be any warning of an attack. Cyber attacks, by their nature covert and fast, invariably involve surprise. The pattern will be hours or days of electronic reconnaissance followed by minutes or seconds of attack. Responding at net speed may not make a difference in these cases. An effective strategy cannot be reactive. To succeed, cybersecurity requires a strategic approach that blends international security actions with domestic measures.

The essential goals for U.S. policy are to associate consequences with malicious cyber actions, reduce misperceptions and miscalculations, and create a framework of rules and expectations for state behavior that constrains the use of cyberattack and other malicious cyber actions. This paper identifies needed changes in U.S. policy to suggest how the United States can successfully engage on cybersecurity with potential military opponents to create a stable international environment. To do this, the United States must make a series of foundational decisions, balancing the costs and benefits of different actions. This is the strategic context for cybersecurity—few if any of the possible responses do not require some trade-off with some other desirable outcome.

The objective of any strategy must be to make cyberspace no riskier for citizens, commerce, or national security than is the case for any other activity. The most difficult obstacle to achieving this is the behavior of states, which have been quick to seize the opportunities of a porous technology and a weakly governed space. States are the source of greatest risk (because they have the greatest capability and resources), and the risk and damage they cause can be managed only by other states. The risk of attack can be lowered by establishing consequences for malicious action or attack through engagement in three interconnected sets of actions:

- Environment shaping actions—understandings among nations about responsible behavior in cyberspace;
- Opponents' shaping actions—reducing misperceptions and miscalculations about the consequences of hostile action against the United States;
- Force-shaping actions—developing the tools and policies for diplomacy, force, and coercion that the United States will use to defend itself and to shape both the international environment and opponents' thinking.

Cybersecurity's Strategic Context

This paper focuses on attack and espionage and makes certain assumptions about vulnerability. It assumes that at a minimum, nations are as vulnerable to cyberattack as they are to attack by any other long-range weapon system. Just as no one could say that a nation could prevent all attacks by aircraft or missiles, one cannot say that all cyberattacks could be blocked. Even if technology were to become perfect (a distant prospect), human error would still allow some probability of success. It assumes that cyber espionage is a dramatic expansion of signals intelligence capabilities and the opportunity for covert acts of political manipulation. The link between espionage and attack is that the techniques for penetration of the target network are, at least initially, largely the same.

Analysis would benefit from a precise definition of "cyberattack." This would limit cyberattack to exploits that cause damage, disruption, or casualties. If we adopt this definition, cyberattack becomes another "weapon" that allows those wielding it to damage or coerce others. By treating cyberattack as just another long-range weapon, we can assess the probable effects of a successful attack, more confidently estimate how nations will plan to use such attacks, and identify the framework of "rules" that should apply to their use.

Only a few nations—Russia, China, Israel, France, the United States, and the United Kingdom—and perhaps a small number of the most sophisticated cyber criminals – currently have the advanced capabilities needed to launch a cyberattack that could do serious and long-term damage equivalent to sabotage or bombing and thus rise to the level of an act of war. A sophisticated attack against infrastructure requires planning, reconnaissance, resources, and skills that are currently available only to these advanced cyber attackers. As part of their larger military planning, these nations have likely planned to launch such attacks in the event of a crisis.[1] Such attacks are not yet within the scope of capabilities possessed by most nonstate hackers.

From a military and intelligence perspective, hostile nations with advanced militaries pose the greatest risk, judging by the capacity for truly damaging cyberattack. Many nations are acquiring military cyber capabilities, but America's most dangerous opponents are China and Russia, which combine advanced capabilities with hostile intent. They exploit U.S. cyber vulnerabilities to degrade U.S. economic, technological, and military "hegemony," and they are the best prepared among likely U.S. opponents for the military use of cyberattack. The examples of their success are too numerous to list. Russia and China have conducted reconnaissance of critical infrastructure to prepare for cyberattack. Pretending that the source of cyber threats is some amorphous and unknowable agglomeration is not supported by data and is not a sound basis for strategy. A review of the publicly known major malicious cyber actions directed against the United States shows that many can be attributed to either Russia or China. America is not in a war; but cyberspace is a contested area. Managing the risks of cyber conflict with Russia and China is of immediate importance for America's national security. Progress in this will help create a more stable environment that will make it easier in the future to deal with erratic states and nonstate actors. For Russia and China, stability is achievable by reducing the disproportionate advantages and relative impunity they currently enjoy.

Other nations hostile to the United States are developing cyberattack capabilities. North Korea has been developing cyber capabilities for more than a decade, but the backwardness of its economy has so far limited its ability to develop cyberattack capabilities. Technological backward-

[1] There is anecdotal evidence that the intelligence services or militaries of several nations have "mapped" digital infrastructure in the United States to provide the capability for cyberattack in the event of a conflict.

ness and political culture remain obstacles to developing strong hacking capabilities; but as with its atomic bomb, North Korea will support sustained investment in developing cyberattack capabilities. North Korea's erratic behavior suggests that it may be willing to use cyberattacks against South Korea, Japan, or U.S. forces in Korea.

Iran has acquired cyberattack capabilities, has a national strategy, created proxy forces, and established military and intelligence entities for cyberattack.[2] It has tested these capabilities against Aramco and a group of leading U.S. banks. Iranian hackers have greater access to the Internet and to the cyber black market than North Korea. Iran, even more than North Korea, could miscalculate the costs of a cyberattack against the United States. Iran sponsors groups like Hezbollah that it has used in the past to attack Americans. The Iranians may believe that these proxies will make it difficult for the United States to attribute an attack, and this could reduce their perception of the risk of a cyberattack on American targets.

In addition to state opponents, anti-American and activist groups have made use of the Internet. As cyberattack capabilities become "commoditized," the temptation for politically motivated groups to use them against vulnerable U.S. targets will increase. We have not seen terrorist groups use cyberattacks—they seem to have neither the capability nor the interest—but since these groups make extensive use of the Internet for political and organizational purposes, they could eventually move to using cyberattacks.

The introduction of new classes of actors has serious implications for stability in cyberspace. Achieving equilibrium—a condition where no party has enough advantage to justify launching an attack—becomes more complicated. This means that the pursuit of stability will be iterative; achieving a stable outcome with major opponents will an initial step, followed when necessary by efforts to deal with new opponents as they gain cyberattack capabilities.

Cybercrime is a tool used by these states, but it is not the greatest source of risk. There is a question of aggregation, and whether the level or cumulative effect of crime and espionage in cyberspace could reach the point where it becomes a threat to U.S. national security. If the amount of cybercrime were to reach a point where legitimate online activity would be greatly curtailed (and we are not at that point), then it would harm the national interest. The cumulative effect of cyber espionage is harder to assess, as many other factors—the ability of the acquirer to use the technology and the rate of technological change (if it takes 10 years for a stolen design to be introduced, the effect may be minimal)—will shape the risk to national security. The greatest risk of cybercrime and cyber espionage comes from the sense of instability they create and from the potential for miscalculation that could lead to escalation from a cyber incident to a more damaging military conflict.

The politics of international cybersecurity are difficult. The United States' potential foreign opponents have disparate views of security and competing goals for international affairs. The end of the Cold War did not mark the end of competition among nations, but this competition often occurs outside the framework of America's expectations about how states will behave. When Russia and China abandoned communism and embraced markets (but not democracy), America expected that they would play by the Western "rules" of international behavior; but in fact they do not regard these rules as binding. We are in an ambiguous period—not great power politics where nations define their policies as competitive with and in opposition to the other—a "zero sum" state

[2] Reuters, "Tehran Sees Cyber Attacks as Greater Threat Than Actual War," *Arab News*, September 26, 2012, http://www.arabnews.com/tehran-sees-cyber-attacks-greater-threat-actual-war; Nasser Karimi, "Iran's Revolutionary Guard Launches Cyber Attack: Report," *Huffington Post*, March 14, 2011, http://www.huffingtonpost.com/2011/03/14/iran-revolutionary-guard-cyber-attack_n_835489.html.

for international affairs, where one nation's gain is another's loss, but also not a single, integrated international community content to accept the United States as its leader.

Besides Russia and China, the last decade has seen the rise of powerful new economies—Brazil, India, and China—and regional powers such as Turkey and South Africa, which increasingly assert regional leadership and seek a larger international role. The assent of these nations is a prerequisite for stability in cyberspace. These new powers are dissatisfied with the international institutions created after 1945 and their deference to a declining Europe. One Brazilian negotiator described this as follows: "Probably those who are most frustrated, and who say they are frustrated, are the Europeans. They think they can still indicate paths which others should follow."[3] The new powers believe to varying degrees that the international order set up after 1945 provides the United States with economic advantages, and that preserving these advantages is the guiding principle for U.S. policy. They want to reduce the U.S. advantage and gain it for themselves. The result has been to create powerful "antibodies," where if the United States is for something, other influential nations are suspicious, if not opposed. The terms of international competition with the emerging powers will also be different—it will be for influence over the structures and rules of global finance and business, rather than over colonies and resources.

This political dynamic is shaped by Western uncertainty and, in Europe, by a degree of remorse. Some call this "strategic timidity." The history of Europe's engagement with the world over the three centuries of Western expansion culminated in horrific struggles that left the continent in ruins and eroded the legitimacy of enlightenment politics. The economic downturn and the possible of Europe's decline into second-class status add to this timidity. The loss of faith among many elites in markets, democratic governments, and the institutions and ideas that have emerged in Europe and North America is one of the greatest weaknesses for the West, particularly as its opponents exploit it.

So far, this situation has reflected a reaction to American power rather than an effort to replace it. There is yet no coherent alternative to the framework of institutions and ideas for international order assembled by the United States and its allies after World War II. None of the new powers offers an alternative vision, only a belief that the current system is inadequate because they were not involved in its creation and that they should play a greater role in its management. The one exception may be China, which has an alternate view of national governance often called "state capitalism" that directly challenges the post–Cold War U.S vision of market democracies. How far China will progress in developing this new construct and how well other nations will receive it remains to be seen. To date, Chinese ideas have been uninspiring.

Eroding U.S. influence will also hamper efforts to create international cybersecurity. Influence is measured by the ability to secure a desired outcome. By this standard, the United States is weaker than it was 12 years ago. Misadventures in the Middle East and a belief in many countries that U.S. policy was largely responsible for the global recession explain this erosion in part, but it also reflects the diffusion of power away from the Atlantic and a declining acceptance of U.S. global leadership. Making networks more secure will make the United States more secure and more powerful, but better cybersecurity will not reverse a decline due to some larger cause.

[3] André Corrêa do Lago, Brazil's chief negotiator at UN Conference on Sustainable Development, cited in Simon Romero and John M. Broder, "Progress on the Sidelines as Rio Conference Ends," *New York Times*, June 23, 2012, http://www.nytimes.com/2012/06/24/world/americas/rio20-conference-ends-with-some-progress-on-the-sidelines.html?_r=1.

Cybersecurity and Grand Strategy

Many of these countries believe that the United States has a grand strategy whose objective is to preserve its military and economic "hegemony," and that it is in their interests to undermine or oppose this hegemonic strategy. Hegemony is a concept introduced by Antonio Gramsci in the 1920s and then adopted by the Soviets in a decades-long propaganda effort to shape international opinion against the United States. This propaganda campaign, like other Soviet propaganda efforts, had measurable success, and the term "hegemon" demonstrates what Fred Ikle called "semantic infiltration," whereby an opponent shapes debate in favorable ways by getting the other side to adopt terms that undermine their own positions. The belief in a U.S. Grand Strategy for continued hegemony in which control of the Internet and information plays a central part shapes foreign reaction to U.S. ideas on cybersecurity.[4]

A grand strategy would describe how America could affect events to create an international order that serves its interests, but the concept of a "grand strategy" is suspect. A "Cyber Grand Strategy" is gibberish. A grand strategy is a response to military and political challenges from other powers. Often, there is an implied commitment to realism, at least among American grand strategists, and a belief that self-interest and competition drive relations among states. Grand strategy then becomes the pragmatic application of the tools of power to serve the national interest.

This sounds quite reasonable, but one great impediment to a new grand strategy is that efforts to design one are increasingly removed from the reality of aggregating resources and applying them to attain specific objectives. Grand strategy has become something of an exercise in rhetoric, to persuade and inspire an audience rather than to plan on how to acquire power or how to choose among competing objectives. America has had only two grand strategies—one for World War II, and one for the Cold War. The immediate conclusion is that a grand strategy makes sense when the United States is in a global war, but at no other time. The World War II strategy identified the ultimate objective (unconditional surrender by the Axis powers), set priorities (Europe before the Pacific), and matched resources to priorities (e.g., the endorsement of strategic bombing that emerged from the Casablanca Conference). It is worth noting that an Army History of Grand Strategy written at the end of World War II concluded that the Grand Strategy

> developed as a product of changing circumstances rather than of a predetermined grand design. Coalition strategy evolved as a result of a complex, continuing process—a constant struggle to adjust ends and means, to reconcile diverse pressures, pulls, and shifting conditions in the global war, and to effect compromises among nations with diverse national interests. That strategy, frequently dictated by necessity, often emerged from events rather than determined them.[5]"

The Cold War Grand strategy avoided direct military conflict between the United States and the Soviets (although there was considerable use of proxies), and it dealt with the issue of linking a new military resource—nuclear weapons—to strategic goals. Development of the strategy occurred in phases—initial analysis in the Truman administration, Kennan's Long Telegram and X Article, and Eisenhower's Project Solarium, with its contending teams of advisers who worked through competing strategic approaches to the Soviet Union. Solarium allowed Eisenhower to lay

[4] See also Joseph S. Nye Jr., *The Future of Power* (New York: Public Affairs, 2011), chap. 6.

[5] Maurice Matloff, gen. ed., *American Military History* (Washington, D.C.: Center of Military History, 1973), http://www.ibiblio.org/hyperwar/AMH/AMH/AMH-21.html.

the foundation for a strategy whose central element lasted for 45 years. The strategy linked containment and deterrence in effective and credible ways that provided a framework for resource allocation, while rejecting more aggressive and confrontational approaches as too risky.

Some in the U.S. national security community in the United Stated call for a new Project Solarium to create a new grand strategy for the United States, and there have been faltering attempts to accomplish this. But Solarium was a product of its times, probably unrepeatable and perhaps unnecessary. World War II and the Cold War posed existential threats to the United States, creating the political will in America to marshal national resources at an extraordinary level in the service of strategy and national defense, creating an unparalleled degree of focus and consensus on international affairs. Whatever threats the United States faces today, including cyberattack, they are not existential and do not generate the same political response. It is this political shortfall more than anything else that makes hollow any effort to create a new grand strategy.

The leaders who created the Cold War Grand Strategy also had a different conceptual framework, shaped by their experiences. They were not international relations theorists but individuals who had learned how to wage a global military conflict in World War II and then were confronted by a powerful and unmistakably hostile authoritarian opponent in a largely military and political conflict. They had not inherited power but had won it, and they did not regard it as an entitlement or as the innate order of things. These factors—searing experience, a single opponent, and a decades-long struggle—allowed for a focus on strategy and strategy making that is unrepeatable today.

The desire for a new Solarium comes after a long period of tentative efforts at to redefine America's international strategy since the Cold War. In some ways, the September 11, 2001, terrorist attacks postponed the debate by providing the United States with a new opponent upon which it could focus its attentions. Terrorism threatens public safety, but not U.S. sovereignty or primacy. By paying less attention to larger strategic challenges (which themselves were not always military), the war on terror damaged U.S. international power and influence. A similar lack of attention to cybersecurity has allowed this to blossom into a major international problem.

Instead of a Grand Strategy, it is more accurate to think of the United States as having long-term interests that it pursues in a consistent (if erratic) fashion. A commitment to open international trade and equal economic opportunities dates to the 19th century. Wilsonian policies added a commitment to democracy and self-determination, and a preference for rules and institutions to shape international affairs rather than force. The end of the World War II saw the addition of a commitment to individual rights. The belief that U.S. interests are best served by creating a stable and prosperous world based on democratic societies where individual rights are respected, where economies are equitably open to trade, and where the rule of law applies to both national and international practices and disputes remains the core of American foreign policy and the best explanatory framework for its actions.

This is not to say that these objectives are consistently pursued by every administration or that other objectives—whether it is the Monroe Doctrine and noninterference in the Western Hemisphere or the global war on terror—do not shape policy and decisions. The long ideological struggle with communism also distorted foreign perceptions of American policy. When compared with some abstract notion of perfection, America falls short. But when compared with the behavior and policies of other nations, it is unique in its pursuit of a just international order.

These core American interests are the objectives that a strategic approach to cybersecurity should pursue. There are tensions, however, between these core goals and cybersecurity. Better cybersecurity requires international cooperation, particularly with Russia and China. America's pursuit of democracy and individual human rights in other nations puts the United States in opposition to these countries' regimes, which are neither democratic nor particularly scrupulous about individual rights. But the United States cannot impose its will on these countries—just as Eisenhower rejected a "roll-back" strategy for confronting the Soviet Union because of its risk and expense, America cannot use force to change China or Russia. Nor are American interests best served by attempting to isolate and contain them (this would fail, in any case, in an interdependent world). Cybersecurity is part of a larger foreign policy problem of defining the bounds of competition and cooperation in a new and multipolar environment.

We can reject statements that America is in a new cold war or a covert cyberwar, as these characterizations are inaccurate and often self-serving. There was a clear demarcation between East and West in the cold War that no longer exists, and U.S. relations with China, Russia, and other nations are today defined by a high degree of openness and deep economic interconnections. These connections do not guarantee peace any more than the connections found before World War I, but they shape policy in ways that are difficult for a nation to manage or control without some degree of cooperation. The defining characteristic of this strategic environment is interdependence, not polarity.

Requirements for a Strategic Framework

America's strategic goals for cyber security should be to reduce impunity and create consequences for malicious action in cyberspace, and create a framework of rules and expectations that reduce misperceptions and miscalculations. These actions would change opponents' calculus to reduce the attractiveness of cyber exploits or attack and reduce the destabilizing effects of cyber exploits on international relations.

The United States will need to balance, as it did for much of the Cold War, the accommodation of potential opponents with a steady pressure to move them in the direction of long-term American interests. This requires steadiness more than confrontation. It also means no compromising of the commitment to democracy or human rights, which risks becoming an implicit quid for better cybersecurity. Should there be a situation where the United States must choose, it should choose its long-term interests. A strategic framework for cybersecuirty should have three central goals: to reduce the chances of miscalculation and misperception by potential attackers; to change opponents calculus of the costs and benefits of cyber espionage and attack; and to build areas of international agreement (such as norms) that promote stability.

We can improve the application of strategy to cyber conflict by recognizing that the behavior of states is largely constant. States are an aggregation of individuals whose actions are shaped by their desires, beliefs, loyalties, and institutions. Over time, it is very possible that the digital revolution will change human behavior in significant ways—this is already observable in how people connect and join groups and even in how they think. But for the most part, a certain constancy can be detected, and the motives and interests of individuals and states are unchanged by the digital revolution.

This constancy of motive and interest is a powerful tool for deconstructing actions and trends in cyberspace. It means that one is talking about a new category of tools and techniques to achieve existing state objectives. And this means that cybersecurity is not much different from any other issue in international security. The same political and economic forces shape it. The specifics of the technology affect both problems and solutions, just as they do in nonproliferation or arms control. There are areas of ambiguity, but cybersecurity is neither sui generis nor subject to such rapid change that intervention is impossible. An emphasis on the speed of technology, and how this limits the scope for government action, is a rhetorical device that can be discarded. Data and precedent allow the United States to test assertions about cyberspace and to more accurately describe problems.

One fundamental decision for strategy is whether cybersecurity is sui generis, requiring new rules and special treatment, or whether it can be fitted into the existing framework of understanding and rules for state behavior. One can reject the arguments that the new cyber environment is unique and that the old rules do not apply. They are historical artifacts of a period of triumphalism, after the Soviet collapse, when it appeared that nations would be replaced by a borderless global society based on market democracies and the rule of law. We mistook, however, the erosion of distance for the erosion of borders and the state. The new technologies make the United States nearer to other nations, not the same. This may be different some decades hence; but for now, cyberspace is not a unique environment, and the normal rules and practices that apply to state behavior (and to the behavior of individuals) can be applied to it.

The reasons for this revolve around the relationship of the state to coercion and legitimacy, and the ability to aggregate and direct resources for use in some strategic purpose. National governments never had a monopoly on force—where the rule of law is weak, there have always been bandits and rebels—but rather a predominant role. This near-monopoly on force continues, and the strength of the institutions that exercise the use of force on behalf of the state has not diminished.

Similarly, predictions about the role of states in cyberspace appear to be derived from initial analyses after the Cold War of the changes that one could expect in international relations. The end of a bipolar world and the widespread adoption of market economies accelerated the creation of a unitary global economy and this implied a similar unitary political structure. It appeared to some observers that in this environment, national government would be less influential and borders less important, perhaps even meaningless. At the same time, a new set of nongovernmental actors would increase their influence and importance. But announcements of the demise of the Westphalian state were premature. States are still the most powerful social institutions, given their centrality in the creation of laws and their ability to aggregate and use force.

A central role for states suggests that it might be useful to reconsider the notion of cyberspace. The term is convenient shorthand for a complex infrastructure—possibly the most complex infrastructure ever assembled—of millions of globally interconnected devices. Like any abstraction, however, it can introduce inaccuracies into our understanding of the terrain and the problems. Cyberspace represents the sum of all devices and social interactions, and gives form to the illusion that it is actually a space, a place, in which we participate in some larger collective experience. Strategies and policy based on an abstraction like "cyberspace" risks compounding inaccuracy.

States are the most powerful actors in cyberspace, given their ability to mobilize and influence the actions of others, to marshal resources, and to harness the tools of force—not just cyber tools but all tools of force and coercion. Borders remain valid; the first airplanes could bypass terrestrial borders with ease until governments developed the technologies to control them. These assertions run counter to the received wisdom upon with Internet policy has been based, many of which date from the 1990s and that decade's optimistic beliefs about globalization, the triumph of market democracies, and the influence of civil society. Cybersecurity is not sui generis but can be incorporated into the existing framework of relations between states and their citizens. This means that the foundation of cybersecurity is agreement among states on how they will use the new technology as a tool of power.

2 THE MILITARY USE OF CYBERATTACK

We need a much more precise understanding of cyberattack and its use as a tool of force and coercion if we are to develop effective strategies. Much of the discussion about the subject we call "cyberwar" reflects a high degree of confusion of what this actually entails—cyberwar appears as a mysterious new form of conflict. But it is not clear that the term "cyberwar" even makes any sense. It may be better to say war, or coercion, or the application of force, using cyber techniques. War uses force to compel an opponent to change in some way they would not otherwise choose to do. Its goals are the destruction of an opponent's capacity and will to resist. Like all weapons except nuclear weapons, cyberattack can contribute to achieving these goals but cannot alone achieve them.

By "cyberattack," we usually mean a software program transmitted over digital networks and installed covertly on a target machine to disrupt data or services or destroy machinery. Similar techniques can also be used for espionage, substituting the exfiltration of data for damage or disruption. Many militaries are developing attack capabilities, but this is not yet some revolutionary and immensely destructive new form of warfare in which any random citizen or hacker can engage at will.

The most likely use of cyberattack will be to identify targets for kinetic attack and "shape the battlefield" in ways that provide advantage against an opponent. Even the most advanced cyberattacks provide limited destructive capability. A cyberattack will not be decisive. Large industrial countries are not easily defeated by a single strike unless it involves nuclear weapons, and cyberattacks do not reach this level of shock and destruction. The destructiveness of cyberattack is overstated. It can cause physical damage to equipment connected to the Internet, but without the shock, confusion, and violence associated with blast damage. The evolving military doctrine for cyberattack involves consideration of how to use cyber techniques and exploits as part of a larger military campaign. In this, the use of cyberattacks intended to cause damage equivalent to that caused by kinetic weapons is best seen as an additional military capability, joining land, sea, and air forces in the combined arms portfolio.

Advanced militaries are embedding cyber capabilities in their existing force structures and military planners are incorporating cyberattack into their doctrines and plans for armed conflict. This is a period of experimentation as militaries try different organizational models and doctrines to see which provide the greatest advantage. Should there be armed conflict, cyberattacks on military forces are certain. Cyber warfare will involve the disruption of crucial network services and data, damage to critical infrastructure, and the creation of uncertainty and doubt among military commanders and political leaders. The best example of this is the 2008 exploit against U.S. military networks, when the Department of Defense's classified SIPRNET was penetrated. No data were exfiltrated, but the implanted malware could have deleted or disrupted stored information to cause immense damage to military command functions.[6]

[6] William Lynn, "Defending a New Domain: The Pentagon's Cyberstrategy," *Foreign Affairs* (September-October 2011), http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain.

Nations and armed groups make decisions on whether to use force based on an internal calculation of the risk and benefits, shaped by their beliefs, knowledge, and cultural predilections. These last three introduce some variation in decision-making, but the fundamental question is whether an attack provides sufficient military or political benefit or advantage to justify the risk of a damaging response or international opprobrium. A potential attacker will weigh the risk of discovery (which many may underestimate), the target's likely response, and the degree to which he cares about the reaction of the international community. The attacker will compare these "costs" against the benefit provided by the damage the attack could produce and by assessing where any disruption caused by cyberattack fits with larger strategic goals.

A campaign limited entirely to cyberattacks would be inadequate to compel or destroy any except the feeblest opponent—cyberattacks do not do enough damage or create enough shock. Cyberattack is not like nuclear war—there are some similarities in decision-making that result from an attacker facing similar decisions on long-range targeting, but these similarities may apply to decisions on the use of any long-range weapon. A better way to think of cyberwar is as a new technique or tool for attack that military commanders will consider among a portfolio of weapons and attacks. Like any weapon, the characteristics of cyberattack make it more advantageous in some situations and less attractive in others. These characteristics will change as attack techniques are refined, making cyberattack less risky to use. The characteristics of cyberattack include high speed, surprise, and covertness. These desirable characteristics must be balanced against cyber's less destructive payload and perhaps "single use"; once a cyber techniques is employed, the target may rapidly develop countermeasures that preclude additional use. These attributes mean that decisions on the use of cyber attack will not differ greatly from a decision to use any other weapon against an opponent—in some circumstances cyber will be an attractive option; in others kinetic weapons may be preferable.

Cyberattacks can be used like both support weapons, to shape the battlefield, and like "strategic" weapons to attack key military and industrial targets in the opponent's homeland.[7] The chief advantage of cyberattack is its effect on the use of information in military operations. Faster decision-making and better information make forces more effective. This "informational advantage" is a legitimate target for counterstrike. By creating uncertainty and doubt in opposing commanders, cyberattacks could cause them to hesitate in making decisions and to act more slowly than would otherwise be the case—the Clausewitzian fog and friction. Cyberattack could provide a tangible military benefit. U.S. forces benefit greatly from their informational advantage—derived from satellites, sensors, and computer networks. Attacking these soft targets would degrade American military capabilities, and they are a certain target for America's opponents in any conflict as opponents seek to gain asymmetric advantage.

Cyberattack reduces the risk of politically damaging images (burning buildings, injured civilians) when compared with a kinetic attack, if care is taken to limit collateral damage through 'weapons" design and precise targeting. Some early cyberattacks (by the United States against Serbia) took down a network, only to discover that innocent civilians also depended on it, or that a target network supported services in a neighboring country. Cyberspace is shaped by business logic, producing a dense set of interconnections that follow commercial rather than geographic boundaries. The environment for cyberattacks is a crowded one, where combatants may be connected to allies, friends,

[7] General James Cartwright (remarks at Schieffer Series: Securing Cyberspace in the 21st Century, CSIS, Washington, D.C., December 6, 2011), http://csis.org/event/schieffer-series-who-commands-commons-securing-cyberspace-21st-century.

and neutral third parties. An attack on a legitimate target may unavoidably damage a neutral party unless it is carefully targeted. "Weapons" design can mitigate the risk of collateral damage, as was the case with Stuxnet, where many systems were infected but only one was damaged. Cyberattack, like aerial bombing, is moving from carpet-bombing to precision as technologies improve.

"Battle damage assessment" before launching a cyberattack—to gauge the effect of digital disruption by mapping out connections and dependencies—will be necessary to make cyberattack consistent with the requirements of international law for armed conflict. As cyberattacks become more precise, the willingness to use them will increase, as it will be easier to manage collateral damage. The context of the attack will also shape decisions for use. In battlefield use, there will be a greater acceptance of collateral damage than there might be in an attack that was part of a covert campaign against a civilian target.

There are parallels between cyberattack and air power. In 1914, the airplane was a contraption of wires and cloth, and the French general who would lead allied forces in World War I dismissively said that it would make a good replacement for a horse. Four years later, aircraft were essential for military operations; and within a decade, every advanced military had some kind of air unit and other militaries sought to acquire them. Cyberattack is on the same path. There are also parallels between cyberattack and ideas on strategic bombing advanced in the 1920s.[8] Just as early air theorists incorrectly believed that the "bomber would always get through," causing war-winning disruption and panic by attacking critical infrastructures, cyber theorists' ascribed almost magical powers to cyberattack.

Attack methods vary in their sophistication, and to a degree, the more sophisticated are also likely to be the more damaging. Sophisticated cyberattacks require an ability to find and exploit technical flaws in an opponent's devices and networks. The most advanced require some degree of engineering expertise. In contrast, "social engineering" relies on tricking a user into opening an email or attached file that then implants malware on a network, or "drive by" downloads, where a visit to a Web site provides an entry point for malicious software. Espionage techniques use similar duplicitous approaches, but face a more difficult task than an attack. To gain intelligence, the opponent must not only penetrate the network but also covertly exfiltrate data. To disrupt, the opponent needs only to implant malware that can interrupt command-and-control networks by damaging the integrity of the information they hold or by erasing it. This requires gaining access to and implanting malicious code on a network, but does not directly produce physical destruction. This lack of physical damage is a source of ambiguity in the application of the laws of armed conflict to cyberwar.

Simpler attacks can include very basic denial-of-service attacks, taking advantage of network configuration errors to deface Web sites (as used against Georgia in 2008) or denial-of-service exploits, essentially flooding the opponents computer networks with incoming traffic, causing them to shut down (as was the case with Estonia in 2007). Neither defacement nor denial-of-service attacks cause damage, although they can have a coercive political effect, and an astute defender can easily block them.

The most sophisticated cyber powers have the ability to disrupt and destroy critical infrastructure systems. The 2011 Stuxnet exploit against Iranian centrifuges and the 2007 Aurora tests at Idaho National Laboratory against electrical generating equipment showed that a high degree of engineering knowledge is required along with an ability to combine a number of advanced intelligence techniques.

[8] For further discussion of the similarities to strategic bombing theory, see James A. Lewis, "Assessing the Risks of Cyber Terrorism, Cyber War and Other Cyber Threats," CSIS, December 2002, http://csis.org/publication/assessing-risks-cyber-terrorism-cyber-war-and-other-cyber-threats.

Serious cyberattack independent of some larger conflict is unlikely. "Pure" cyberwar—"keyboard versus keyboard" or "geek versus geek"—is unlikely. Cyberattacks are fast, cheap, and moderately destructive, but no one would plan to fight using only cyber weapons. They are not destructive enough to damage an opponent's will and capacity to resist. Cyberattacks will not defeat a large and powerful opponent. In most scenarios involving state actors, these attacks will not be "cyber only" incidents unless an opponent makes major miscalculations or unless their tolerance for risk is high, and risk tolerance may be a function of their belief that the action is covert (and thus less likely to trigger a damaging response).

One assumption that must be carefully avoided is that an attacker will follow U.S. organization and doctrine. At an operational level, one should expect cyberattacks against deployed and supporting military forces to combine electronic warfare (EW) and (most likely) anti-satellite attacks (ASAT). If the opponent's goal is to degrade commander confidence and reduce operational efficiency and the informational advantage of American forces, an astute attacker of even moderate size will combine cyber, EW, and ASAT attacks on U.S. assets. Given the interdependence of weapons, sensors, and networks, an attack on these informational assets may be as effective as eliminating the weapons themselves.

Russia's February 2010 "New Military Doctrine" recognized information warfare as a tool for conducting psychological warfare and shaping public opinion.[9] The doctrine promotes information warfare as a force multiplier to weaken the command-and-control capability of the enemy. Information attacks would be deployed in two stages—in peacetime and during battle. Information warfare in peacetime involves preparation for conflict by ensuring the security of the government, by strengthening defenses and through deceptive measures designed to shield Russian forces and targets. During battle, Russia would use information warfare to dominate the enemy at the strategic, operational, and tactical levels.

Public discussion of Russian thinking on information warfare has four components. First, it involves the destruction of the opponent's command-and-control centers. Second, it includes signals intelligence and the interception of enemy communications. Third, enemy information systems are accessed and sabotaged to disrupt normal operations. The last component of information warfare is psychological, and entails shaping the opinions and views of the enemy. Russian doctrine requires pre-strike planning, where the targets are identified, along with opinion shaping campaigns before the onset of kinetic conflict.

China's cyber strategy focuses on asymmetric warfare and the necessity of quickly disarming opponents through a debilitating first strike. The enemy would suffer a debilitating first strike in the opening stages of conflict that would limit its ability to retaliate. The main doctrine on cyberwar strategy advocates for a combination of cyber warfare and electronic warfare capabilities in the early stages of conflict to paralyze command-and-control and intelligence centers. China will combine cyber warfare, electronic warfare, and ASAT techniques to achieve this goal against what the People's Liberation Army sees as its most likely opponent, the United States. This is in addition to the "technical reconnaissance" (or espionage) involving the collection, analysis, and exploitation of electronic information.

In facing these skilled opponents, it would be easy to leap to the conclusion that offensive cyber operations always have an advantage over a defender. This deserves some scrutiny. In the

[9] Russia does not use the term "cyberattack"; it sees "cyber" as a technical subset of "information warfare." See Paul Cornish et al., *On Cyber Warfare*, a Chatham House report (London: Chatham House, November 2010), http://www.chathamhouse.org/sites/default/files/public/Research/International%20Security/r1110_cyberwarfare.pdf.

1920s and 1930s, there was a school of thought that said “the bomber will always get through,” a part of the larger overestimation of the utility and effect of strategic bombing. Similar conditions apply to cyberattack. It is true, given the speed at which a cyberattack may occur, that absent forewarning, the probability of interception is very low. Cyberattack is essentially a surprise attack, and thus much of the literature on surprise attacks, starting with Wohlstetter’s classic study, applies to it.[10] If there is no forewarning, and little preparation, cyberattack may be unpreventable. This “surprise’ aspect has serious implications for policy.

First, the emphasis on situational awareness in the public discussion reflects the desire for forewarning and a hope that it will be possible to block attacks at “net speed.” If the attack vectors are know to a defender, a net speed defense is possible. For some malicious cyber activities, such as crime, the techniques are used against multiple targets over a period of time, meaning that if the initial exploits are detected it may be possible to prevent additional use. The same logic may not apply to cyberattacks intended to cause physical damage. The attacker will have attempted to avoid forewarning and may accelerate any action to defeat a response. Whereas espionage may involve sitting on a network for months, an attack may begin and end within a few seconds. The nature of cyberattack is such that there is a limited chance of stopping an attack once it is launched. There will not be any warning.

Active intelligence collection against a potential opponent on its plans and attack capabilities will provide opportunities to block or preempt an attack. One decision for the defender is where to take action—on computers and networks within its own national jurisdiction, or on networks subject to the opponent’s jurisdiction. An action on networks within its own jurisdiction is politically less risky, but it could be hazardous to let an attack reach that point—there would be only a few seconds in which to react. A preemptive attack returns control over the pace of engagement to the defender but brings the increased political risk incumbent on attacking another nation.

The threat of a preemptive strike could have deterrent effect if the opponent believes that it is likely to be detected in advance, but not if it believes its actions are covert. The speed at which cyberattack occurs and the emphasis on covertness mean that increased attention to indicators and warning of a surprise attack are essential. Defensive planning will need to assume the likelihood of successful attack, and consider how to increase resilience at a national level. Response to cyberattack, while important, does not have to be at net speed. It took the United States four months to send Doolittle in response Pearl Harbor.

The inadequacies of defense against cyberattack may be reduced in the future, but not to the point where cyberattack or cyber espionage will be unattractive options for a nation to pursue. Digital capabilities for espionage and attack create a new class of military problem that must form part of any serious strategy for national defense. This strategy must determine what can be done in advance of any conflict to shape, constrain, and perhaps deter cyberattack.

Framing Cyberattack with International Law

The starting point for a discussion of the applicability of international law to cyberattacks is to accept that if it is legitimate to attack a target physically, it is also legitimate to attack it using cyber weapons. The United States would benefit from clear international understandings that accepted the application of international law to shape both decision-making on the use of cyber attacks and on cyber

[10] Roberta Wohlstetter, *Pearl Harbor: Warning and Decision* (Stanford, Calif.: Stanford University Press, 1962).

“weapons” design. While there is some disagreement as to whether one can apply the existing legal framework for warfare to cyber conflict or whether a new legal framework is needed, a review of the applicability of existing law of armed conflict suggests that if one approaches cyberattack as another new military technology, current international law can be applied to it. Some issues involving sovereignty, the nature of combatants, and the use of force may need to be expanded or redefined.

It could be argued that these laws are honored more in their breach than by observance, and that certain classes of non-Western or nonstate opponents feel little bound by them. We cannot dismiss these laws, however, for three reasons. First, an important U.S. goal is to expand adherence to the rule of law. Second, being seen as acting in accordance with international law (or being able to argue that one's actions are in accordance with it) provides substantial benefit in international politics and public opinion. Third and finally, U.S. policymaking, for better or for worse, is heavily legalistic and decisions must be grounded in some way in the national and international legal framework.

Two separate bodies of law apply to warfare: “jus ad bellum,” the legal framework for a decision to resort to the use of force, and “jus in bello,” the laws governing the conduct of hostilities. Jus ad bellum guides political leaders' decision as to whether an incident justifies an armed response under the provisions of the United Nations Charter on a country's right to use force in self-defense. These laws are derived from international conventions and treaties (e.g., The Hague Convention and Geneva Convention) and from customary international law. They set forth rules that govern the use of force during armed conflict. Three principles from the Law of War in particular establish a framework for judging the legality of using different forms of cyberattack during an armed conflict: distinction, proportionality, and discrimination.

A short summary of these laws shows their applicability to cyber conflict. Distinction requires attacks to be limited to legitimate military objectives. Article 23 of The Hague Convention, for example, forbids belligerents “to destroy or seize the enemy's property, unless such destruction or seizure be imperatively demanded by the necessities of war.” Proportionality requires that the use of force must be limited to that which is necessary to meet an imminent or actual armed attack and must be proportionate to the threat. Attacks on a military objective that cause incidental loss of life or injury to civilians or damage civilian property in excess of that needed to obtain concrete and direct military advantage should be avoided. Discrimination calls for attacks to be limited to a specific military objective and to avoid attacks that are indiscriminate or haphazard in their inclusion of civilian targets. Belligerents are to take all necessary steps to avoid damage to “buildings dedicated to religion, art, science, or charitable purposes, historic monuments, hospitals, and places where the sick and wounded are collected, provided they are not being used at the time for military purposes.”

Discussion of making some set of networks “sanctuaries” that participants agree not to attack ignores both precedent and the degree of interconnection between legitimate and “illegitimate” targets. Existing international law would seem to prohibit attacks on purely civilian infrastructure when the resultant disruption or destruction would not produce meaningful military advantage (recognizing that there are no objective standards for determining what constitutes meaningful military advantage, these decisions are left to military commanders and a nation's political leadership).

Additionally, these principles imply that an attacker would need to assess the potential for collateral damage to civilian targets for a cyberattack to be lawful. To be consistent with the laws of war, the use of cyberattacks during conflict would face the same constraints as attacks using kinetic weapons. The goal of the protections for civilians found in the laws of war is not to shield them from the dangers of military operations but to avoid capricious attacks undertaken solely to harm civilian targets.

Deciding when an attack is consistent with international law requires a commander to ascertain that overriding military necessity justifies the action. Cyberattacks against critical infrastructure do not allow for the possibility of sanctuary, although it might be possible to rule out direct attacks against targets like hospitals or churches unless they were being used for military purposes. A decision to disrupt power supplies is legal under the existing rule for armed conflict and is a normal feature of modern warfare, but will affect both military and civilian targets.

No target is excluded from attack if the resultant disruption or destruction would produce meaningful military advantage. The decision as to what constitutes meaningful military advantage is left to military commanders and a nation's political leadership. Military forces can attack a hospital if an opponent is firing from within it. Civilian infrastructure is routinely targeted, and its destruction is a core tenet of air power theories. The goal of the protections found in the laws of war is to not to make civilians invulnerable to the dangers of military operations but to avoid capricious attacks undertaken solely to harm innocent targets.

The right of self-defense applies whether the attacker is a state or nonstate actor, so the official or unofficial nature of the attack is not a helpful distinction for determining whether a military response is appropriate. The legal framework for conflict applies to both state and nonstate actors, but the decision as to when to apply it depends in large part on whether an action is deemed to have involved the use of force (e.g., it is an armed attack). Terrorists are not "lawful combatants" under international law, nor do they usually feel that such laws constrain them.

The applicability of international law regarding the use of force and the right to self-defense will remain somewhat ambiguous, however, until it is possible to clarify when and under what circumstances, a disruptive exploit in cyberspace could be considered an armed attack. The United States would benefit from clarifying first for itself and then internationally what kinds of cyber actions triggered the right to self-defense and justify and armed response. Determining the threshold that distinguishes cyberattack from other malicious activity is crucial for policy.

The 2007 Estonian episode points to an area of ambiguity in the application of existing law. It is possible to imagine circumstances where a malicious cyber action could threaten territorial integrity or political independence without physical destruction. One fundamental question is whether a cyber exploit must produce physical damage and casualties to be regarded as an act of force, or whether other, intangible damage inflicted outside armed conflict can also be considered a use of force and an act of war. Jean Pictet, a Swiss jurist who was the primary author of the Commentaries on the Geneva Conventions that guide the application of the laws of armed conflict, identified three criteria for applicability: scope, duration, and intensity.[11] These criteria, along with the idea of equivalence, provide additional measures for deciding when a malicious action qualifies as a cyberattack.

The 2007 cyber actions against Estonia provide a useful test for these thresholds. Some people say that "cyberattacks" brought Estonia to its knees. This is nonsense; the effect of the "distributed denial-of-service attacks" was limited, and the Estonians showed remarkable resilience in responding to them. The attacks, however, put the Estonian government under intense political pressure, in part because of uncertainty as to whether the cyber actions were a precursor to a Russian invasion or a more damaging cyberattack. The "denial-of-service" actions were a politically coercive action, but it was not war or an attack. This is why NATO did not invoke its Article 5 obligations (collective defense) to assist Estonia. If the denial-of-service exploits against Estonia had been extended in time, if the scope

[11] International Committee of the Red Cross, "Commentary on the Geneva Conventions of 12 August 1949."

of the denial-of-service attack and its effect had been equivalent to a naval blockade, then the actions could have justified a military response even in the absence of physical damage or casualties.

Similar exploits against Georgia in 2008 were reportedly synchronized with Russian military activities, implying a high degree of coordination with the Russian government. It was for political purposes, and it was what the Russians call information warfare. It was to have a political effect; it increased the pressure on the Georgian government. Nevertheless, the intent of these exploits was not to damage or destroy infrastructure or disrupt critical services. The Georgian incidents are a useful indicator of Russian military doctrine and how Russia may use cyber techniques in future conflicts to shape opponent decisions; but by themselves, they do not qualify as an attack. In contrast, Stuxnet did cause damage, could be considered the use of force, and presented Iran with the decision (pace the evidentiary problems of attribution) whether to regard it as a casus belli.

Deciding whether a cyberattack triggered the right to use force in self-defense (or alliance commitments for mutual defense) could be based on the effect of the attack. A malicious action in cyberspace that produced an effect that is equivalent to a physical attack, or an exploit that caused substantial death or physical destruction similar to the use of a weapon like bombs or missiles, would be an armed attack. To trigger the right of self-defense, a nation's authorities would need to decide if a cyber exploit threatened its territorial integrity or political independence. Most malicious activity in cyberspace does not, under existing precedent, meet this threshold. A cyber exploit that was a violation of sovereignty is by itself not sufficient. Crime and espionage are not considered to be the use of force or an attack. Suggestions have been made that the United States would benefit if it asserted a lower threshold for the use of military force in response to a malicious cyber act, but the effect of such a change could be destabilizing and would harm U.S. interests.

We can begin identifying this threshold if we accept that it is the line between reconnaissance or exploitation (espionage and crime) and disruption and damage. We could consider crossing this line as escalating any cyber conflict. A secondary threshold could be the move from disrupting or damaging military targets to damaging critical infrastructure or other civilian targets. Norms or thresholds like these would accept a certain level of conflict in cyberspace but would attempt to increase the risk and consequences for states if they (or actors in their countries) move beyond espionage and crime.

One can distinguish between a criminal act and the use of force in several ways. Criminal acts rarely threaten "the territorial integrity or political independence" of a state. Espionage (in the sense of the illicit acquisition of information by agents of another state) is a crime under every nation's laws, but it is an exaggeration to say that it threatens territorial integrity or political independence. Espionage is not the same as covert political actions undertaken by intelligence agencies to shape or destabilize the political situation in another country to create coercive pressures on that state, Estonia being the foremost example of this kind of behavior.

Clearer understandings of the nature of sovereignty in cyberspace and the development of shared understandings among states could also reduce risk of damage or conflict escalation. While the discussion of sovereignty in cyberspace has been confused, and the notion that it is a "commons" is unhelpful[12]—cyberspace is much more like a condominium, where multiple owners

[12] White House, "Cyberspace Policy Review," May 2009, B4, http://www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf. For a comprehensive discussion of the "Commons," see Michèle A. Flournoy and Shawn Brimley, eds., *Finding Our Way: Debating American Grand Strategy* (Washington, D.C.: Center for a New American Security, June 2008), http://www.cnas.org/files/documents/publications/FlournoyBrimley_Finding%20Our%20Way_June08.pdf.

share control—there is at a minimum no disagreement over sovereignty as it applies to the victim. If the victim determines that an attacker is another state, the sovereignty of the target nation has been violated and a response could involve retaliatory military action. If an intruder is a private individual (not acting at the behest of a foreign government), the intrusion is a crime, not a violation of sovereignty. The dilemma, of course, is that it can be difficult to determine if the attacker is a state or a private party, and thus whether sovereignty has been violated.

We can reduce uncertainty by defining the probable outcomes of cyber actions more carefully. There are gray areas, of course, particularly when this disruption rises to the level of the use of force. A denial-of-service attack such as those launched against Estonia would not be considered an act of war unless it was extensive and prolonged, having essentially the same effect as a naval blockade on the target country's commerce. We can refine strategy by distinguishing between actions in cyberspace that can be considered as the use of force and those that are not. The distinction is important because it may be easier to find agreement on the use of force rather than on whose intent is espionage.

Expanding the definition of cyberattack to include access to information makes little sense. It distorts long-standing ideas on warfare and military action by disconnecting them from the concept of the use of armed force and violence. The use of force produces physical harm. It is central to defining attack and warfare. The concept of force is derived from the UN Charter and The Hague Convention and Geneva Convention. Publishing or sharing an idea is not the use of force. While an expanded definition of warfare may serve the political interests of authoritarian regimes, it is not an accurate description of military action or attack.

Discussions of cyberattack are also complicated by what can be termed the "overflight" issue. Almost all attacks require traversing third-party networks to reach the target. Cyberattacks are covert or clandestine, in that they pose as legitimate commercial traffic that is permitted to cross frontiers under existing commercial law and "interconnect" agreements among Tier 1 service providers.[13] Currently, this issue can be disregarded as the practice among major telecommunications companies is to simply pass traffic without review, but as nations increase their capability to monitor what crosses their networks (perhaps using a technology known as "deep packet inspection"), in order to intercept malicious traffic aimed at themselves, it may become harder to plead ignorance.

Assigning or limiting the responsibility of states to discover malicious traffic aimed at someone else is an area where physical precedent may not be useful. If a state chooses to let "hostile code" into its territory, it is witting and may become a party to the conflict. Few states now admit knowledge as to what passes over their networks en route to somewhere else, or what the intent of that traffic may be, but over time this ignorance will shrink unless willfully asserted. Should deep packet inspection capabilities become widespread, understandings on passage rights or restrictions on military traffic (one approach would be akin to the Montreux Convention; another would be to continue to let states remain unwitting and rely on commercial contracts provide unrestricted passage) may become essential.

Russia has argued that political actions are in fact the core of the new kind of warfare and the issue is really "information warfare" rather than "cyber warfare." Russian officials have said that information is a weapon and that the United States uses information to destabilize governments it opposes. Information is a threat to authoritarian regimes, and they want to limit access to Web sites and social networks. This is more of a political ploy than an area of ambiguity in the application of international law to cyberattack.

[13] CSIS, "Cyber Security for the 44th Presidency: Telecommunications Task Group Final Report," October 2008, http://csis.org/files/media/csis/pubs/081028_telecomm_task_group.pdf.

This suggests that in most cases, existing laws for armed conflict can be applied to cyberattack, but that there are areas of ambiguity involving the violation of third-party sovereignty, the use of cyberattacks by terrorists, and the amount and nature of damage from cyberattack that could be interpreted as an act of war. Some operational issues, such as the amount of prior assessment of collateral damage required to make an attack consistent with the laws of war, are also unclear. There are yet few precedents for resolving these ambiguities, but it may be possible to clarify them as we gain experience, through international discussions or through further analyses and exercises.

The dilemma in their application is that it seems likely that in most cases, cyberattack will not be used as the equivalent of a kinetic weapon (e.g., to cause physical destruction). The legal framework for conflict is largely designed for tangible effects; cyberattacks will often create intangible damage. They will be used to shape the battlefield in advantageous ways by manipulating information to shape opponents' perceptions and decisions. Trickery, ruses, and stratagems have been part of the commander's portfolio since the first combat. Genghis Khan pretended to retreat to lure opponents into an ambush. Eisenhower used inflatable tanks and radio signals that mimicked a huge force to persuade the Germans that he would invade somewhere other than Normandy. Cyberspace creates a new dimension for this age-old military technique and for the larger policy discussion that began with the widespread adoption of the Internet on how we apply rules designed for physical actions to cyber events.

Scenarios for Cyberattack

Cyberattacks have both tactical and strategic applications. They can be used as a support weapon, useful to shape the battlefield in advantageous ways, or like "strategic" weapons, to attack key military and industrial targets in the opponent's homeland. Unlike traditional "strategic weapons (e.g., nuclear) their effect is neither massively destructive nor fatal; nor do they pose an existential threat to nations. Cyberattack can be compared with a missile, offering fast, long-range strike, with greater covertness (perhaps) and a smaller destructive payload. This limited destructive capability does not mean that we should welcome the disruption of an artificial financial panic or a blackout that could last weeks, but the destructive power of cyberattack is nowhere near that of nuclear weapons or even a sustained assault using kinetic weapons.[14]

Cyber techniques are a recognized military and intelligence capability that have been in use for years. Perhaps 40 nations have or are acquiring the ability to use cyber techniques in combat. Most of these national programs are shrouded in secrecy, but cyberattack is an accepted part of the portfolio of every advanced military and a capability many wish to acquire. Political leaders and military commanders have a new tool for attack. It has certain unique advantages when compared with other modes of attack, but also some inherent limitations. Three scenarios for military use define the problem for developing strategy.

The most probable uses of cyberattack will be to "shape the battlefield," to degrade opponent sensors networks and decision-making. A plausible tactical use of cyberattack would be to disrupt air defense networks. "Networked" air defense are many times more effective than non-networked air defenses. These networks use computers to link radar systems, missiles, and guns into a mesh that is much more difficult for aircraft to penetrate.

[14] An example of preparation for this kind of major engagement might be found in DARPA's "Plan X" program, https://www.fbo.gov/utils/view?id=f69bba51a9047620f2e5c3a6857e6f6b.

Some air defense networks use dedicated communications grids to reduce the risk of penetration, but there are still many avenues for entry. The most obvious is through radar systems, which are essentially radio transmitters and receivers, and anything that receives radio transmission is susceptible to penetration. Radar sends out a signal that bounces of objects like aircraft. The returning signal is received by the radar and then processed by a computer. If this computer is networked with other air defense sites and equipment, an astute attacker could transmit their own radio message on the same frequency that the radar uses to introduce a poisoned signal into the receiving computer and then into the entire network. This would allow the attacker to obscure what the radar is detecting, reorient sensors and weapons away from an attack, and see what the opponent radar operator is seeing on their screen.

This approach blends traditional electronic warfare and its jamming and exploitation of opponent signals with cyberattack and the disruption of opponents' software programs and computer systems. The poisoned radio signal can be transmitted from an aircraft, an unmanned aerial vehicle, or a specialized missile. Opponents' radars are identified and then are sent data that allows them to be controlled, have their radars redirected, or false targets created. For example, there has been some public discussion of how Israel, when it launched air attacks in September 2007 against an alleged Syrian nuclear facility, first used blended electronic warfare and cyberattacks to disable Syria's air defense network, which, like many modern air defenses, relies on computer networks. One story has it that the Israelis changed the computer programs to make it seem as if Syrian radars were detecting nothing, even as Israeli jets entered and crossed Syrian territory.[15]

This kind of cyberattack poses little risk of collateral damage since it is focused on military targets. Disruption of air defenses has been a normal part of the planning for any air campaign since the 1940s, so the addition of a cyber aspect is certain. To the extent that land or naval forces rely on sensors, weapons, and commanders linked by computer networks, similar cyberattacks could be launched against them. Logistics systems are a likely target. Cyberattacks could aim to reroute supplies to the wrong location, send the wrong items, or disrupt delivery schedules. These sorts of incidents are already part of the normal "friction" of war. Computers have helped reduce this, but offer an attacker the opportunity to amplify this friction in ways that damage opponent operations.

Naval operations are vulnerable to cyberattack, relying on networks of sensors and weapons ("cooperative engagement capability," developed in the 1980s, is an early example). Attacks on naval network will also blend cyberwar and EW, but could entail attacks on space assets as well. Satellite for positioning, communications, and surveillance are integral to American naval operations. A sudden strike that disabled the warning and communications functions immediately before a high-speed missile attack on an aircraft carrier would improve the chances of success. An opponent could attempt to disable satellite assets that support naval operations through a cyberattack on the satellite control (or the satellite itself), through jamming or through kinetic attack. The United States, China, and Russia have tested all these types of attacks. Many other countries have tested jamming and cyberattacks on satellites and their controlling ground stations. Cyberattacks on space assets are also certain in conflicts involving naval or air forces, which rely heavily on space assets.

[15] See David A. Fulghum, "Why Syria's Air Defenses Failed to Detect Israelis," *Aviation Week*, October 3, 2007, http://www.aviationweek.com/Blogs.aspx?plckBlogId=Blog:27ec4a53-dcc8-42d0-bd3a-01329aef79a7&plckController=Blog&plckScript=blogScript&plckElementId=blogDest&plckBlogPage=BlogViewPost&plckPostId=Blog%253a27ec4a53-dcc8-42d0-bd3a-01329aef79a7Post%253a2710d024-5eda-416c-b117-ae6d649146cd.

For exploits against satellites, the political calculation an attacker would need to make becomes more complicated. Attacks on air defense networks are a routine part of planning and are consistent with the laws of armed conflict. Attacks on naval forces that blended anti-satellite attacks considerably increase political risk by extending the conflict into space, particularly if any kind of kinetic anti-satellite weapons is used. It may be necessary to degrade or damage third-party commercial satellites, complicating the politics of an attack. But a "pure" cyberattack on naval forces that did not include attacks on space assets would have a lesser chance of success (even if it blended cyberwar and EW). Opponents will probably look for ways to argue that their attacks are proportional and launched against purely military targets (military satellites carrying sensors rather than civilian communications satellites) with little risk of collateral damage to third-party or civilian assets, and this may influence in some way their decisions on tactics and the use of cyberattacks.

Similar calculations will shape decisions and doctrine on the use of cyberattacks against civilian targets. The fundamental strategic calculation is whether the military benefits of a cyberattack (by itself, or jointly with electronic warfare and ASAT) outweigh the political risk. Political risk has two components: the reaction of the international community and the reaction of the opponent.

Opprobrium or condemnation by the international community has some shaping influence. The political reaction of the opponent nation is much more important, although it is often underestimated by military planners. A surprise attack will be perceived as "unfair" and provoke hostility. Attacks on civilian targets, based on studies of target populations since the 1940s, are as likely to stiffen resolve or continue the fight to exact revenge.[16] Attacks on civilian targets, particularly in the opponent's homeland, could be seen as an unacceptable broadening of the conflict. The emotional reaction to attacks on civilian targets could escalate the scope and intensity of the conflict. The same is not true for attacks on military networks; these should be expected as a normal element of warfare. Related attacks including military space assets in combination with cyberattack are less likely, as they increase political risk, but cannot be ruled out. Attacks on civilian assets (essentially the application of strategic bombing principles to cyberattack) would unavoidably be seen as a major escalation of the conflict.

Serious cyberattack independent of some larger conflict is unlikely. To transpose cyber to the physical world, there are remarkably few instances of a nation engaging in covert sabotage attacks against another nation (particularly larger powers) unless they were seeking to provoke or if conflict was imminent. The political threshold for serious cyberattack (as opposed to espionage) by a nation-state is very high, likely as high as the threshold for conventional military action. At a minimum, this suggests that a serious cyberattack is a precursor, a warning, that some more serious conflict is about to begin.

Absent such larger conflict, however, a nation-state is no more likely to launch a serious cyberattack than it is to shoot a random missile at an opponent.[17] The risk is too great and the benefits of a cyberattack by itself are too small for political leaders to authorize the use of this capability in anything short of a situation where they had already decided on military action. Cyber weapons are not decisive; cyberattack by itself will not win a conflict, particularly against a large and powerful opponent. It is striking that, to date; no cyber "attack" that rises above the level of espionage or crime has been launched outside of a military conflict.

[16] U.S. Strategic Bombing Survey, "Summary Report (European War)," 1945, http://www.anesi.com/ussbs02.htm.

[17] This raises the question of the command and control of cyber capabilities and the possibility of an inadvertent attack.

In fact, there may be an implicit escalatory ladder for cyberattacks that ranges from political coercion to physical destruction. Escalation is also determined by the nature and location of the target, ranging from deployed military forces in the combat zone to civilian targets in the opponent homeland. Nations may intuitively recognize that some cyber actions create greater risk than others and shape their strategies accordingly.

Even in a conflict, a decision to strike civilian targets in an opponent's homeland using cyber weapons is a major step that brings the risk of serious escalation. Engaging American military forces overseas is not the same as attacking critical infrastructure in the United States. Nations may reserve these serious cyberattacks against targets in the opponent's homeland for either retaliation for attacks against their own homeland or for when they are in extremis.

Nuclear doctrine very quickly settled into a pattern where the use of nuclear attack was reserved, to be used only in extremis.[18] Careful calculation is needed to see if the same should be true for cyberattacks against civilian targets in an opponent's sovereign territory. The United States and its opponents are unlikely to renounce the use of some forms of cyber action involving espionage or low-level disruption, but a more serious attack may likely be reserved for the most serious situations. The deterrent value of a cyber response might be increased if U.S. public doctrine reserved the most serious attacks (those that led to a major disruption of critical infrastructure) for retaliation.

Some attackers may not be deterred by any of these political risks. They may overestimate the destructiveness of a cyberattack and gamble that it will provide unassailable advantage or that the conflict will be of such short duration that the temporary disruptions of cyberattacks will have an appreciable effect. They may underestimate the response of the opponent government or population. They may be more tolerant of risk (particularly if they are a nonstate actor). During in extremis situations, when the attacker believed themselves on the brink of defeat, they may undertake risky attacks to either gain time or gain advantage in talks to end the conflict. Miscalculation based on misperception is natural, and experience suggests that an attacker will begin by overestimating the benefits and underestimating the costs.

Cyber techniques can also be used to mislead an opponent, particularly in conflict with insurgents. We do not fully appreciate how much these groups rely on commercially available information technology. Being able to penetrate the communications networks that an insurgent group uses for command and control would allow you to direct them to an ambush site or to fire on themselves. As one general officer said, "I can tell you that as a commander in Afghanistan in the year 2010, I was able to use my cyber operations against my adversary with great impact. I was able to get inside his nets, infect his command and control, and in fact defend myself against his almost constant incursions to get inside my wire, to affect my operations."[19]

One technique is to track the personal mobile devices most combatants carry to determine their location and their movements. A company or brigade moving north could be tracked by following the mobile device signals. This applies to both insurgents and conventional forces. It is possible, however, to modify the information on the location and movement of the personal device to give misleading information. Movement could be concealed and opponents duped. Cyberspace creates a new dimension for manipulation of opponents' perceptions and planning that can shape the battlefield in advantageous ways.

[18] This idea comes from Arnold Kanter, former U.S. undersecretary of state for political affairs.

[19] Remarks by Lieutenant General Richard Mills (USMC), AFCEA, August 15, 2012.

Penetration and corruption of guidance system data will undoubtedly become a primary focus for military cyberattacks. Most advanced weapons now depend on software to perform. In 2007, a software error caused the computers that provided navigation, altitude awareness, and fuel status crashed when six F-22s crossed the international dateline. The F-22s turned back and were forced to rely on other aircraft for navigation. An earlier generation of aircraft, the F-16, is inherently unstable and cannot be flown without computer control. A reasonable goal for an opponent would be to attempt to introduce a similar effect in flight computers. Being able to access and corrupt that weapon's software before or during battle could significantly degrade weapons performance. If software malfunctions, weapons do not work.[20]

For example, a network penetration could surreptitiously tamper with missile guidance software to create launch failure. To an observer, it would appear that the failure was normal. Critical software in precision munitions guidance could be changed to degrade accuracy. Weapons increasingly use commercial software (to save money), introducing supply chain vulnerabilities by providing potential attacker with an opportunity to better understand how the software is written and used. Even tiny changes can greatly affect a weapon's performance.

Access to the weapon's software to disrupt it could take place during development (supply chain contamination), while the weapon is connected to a network (for monitoring or mission programming), or through a signal transmitted to the weapons system, for example, a radar sends and receives radio signals. This provides an entry point, even if the radar (or the weapons system carrying the radar) is not attached to the Internet or some other public network. Radars receive a signal, process it, and then pass it through a dedicated network to another system, operator, or weapon. A false signal could introduce malicious code to either the processing or the network through this channel. These attacks will blend electronic and cyber warfare techniques to gain entry and then to disrupt. Tampering with opponents' weapons and sensors will probably be a more important concern for militaries than attacking critical civilian infrastructure.

Degradation of an opponent's military networks, information, or critical infrastructure is a legitimate military objective, but the degree of interconnection with third parties complicates the use of cyber weapons. An opponent that exploits interconnectivity and relies on third-party commercial service will be able to complicate an attack and perhaps constrain an attacker. Combatants and noncombatants may even depend on each other in cyberspace in ways that mean an attack on a legitimate target may unavoidably damage a neutral party. This means that the political consequences of some forms of cyberattack are greater and require greater attention from the political leadership before action is authorized.

Strategic Use

Popular accounts of cyberattack often talk about "Pearl Harbors" or even "Armageddons" that cyberattacks will eventually produce. Let the United States dismiss these, with one exception. Nonstate actors have the ability to disrupt computer networks, most effectively by inserting malware that could erase data or by locking users out of essential computing and network services. This could produce blackouts, traffic jams, and business disruptions—all relatively temporary and all with little consequences for national security. These kinds of cyber incidents (they are not really

[20] Major Dani Johnson (USAF), "Lockheed's F-22 Raptor Gets Zapped by International Date Line: Raptors arrive at Kadena," February 26, 2007, http://www.af.mil/news/story.asp?storyID=123041567.

attacks, since they fail to produce damage or casualties) would not degrade America's military capabilities or its long-term economic strength.

Nor are they likely to be connected to some strategic purpose, although jihadis or anarchists may be attracted to straightforward disruption. The fact that these incidents have little strategic consequence does not mean that America should not try to prevent them or that it should continue its current posture of having no adequate defense against such disruptions, but these groups do not yet have the capability to create the levels of damage and destruction found in "weapons-grade" malware. For the countries that do have this destructive capability, they could consider cyberattacks, with their speed, range, and covertness, a useful supplement to existing "strategic" strike capabilities.

In the Cold War, the United States had a Single Integrated Operations Plan (SIOP) for the simultaneous use of nuclear weapons. Using satellite reconnaissance and other intelligence techniques, the United States identified hundreds of military, industrial, and leadership targets in the Soviet Union and China and allocated nuclear weapons to destroy each target. The plan was regularly modified and updated and provided options that ranged from near total destruction of the opponent's military and industrial capability (this was in an era when the United States had thousands of warheads) to more "surgical" interventions targeting specific facilities or regions.

As an aside, informal discussions with Chinese military officers suggest that the Chinese underestimated the effect of nuclear weapons and overestimated their ability to survive. At the end of the Cold War, the maximum set of strikes against industrial and military targets (which would have required several hundred nuclear weapons) would have killed two-thirds of China's population and destroyed most of its industry. Moving industrial centers inland makes no difference to a warhead reentering from space. It is not an exact comparison, but it raises the question of whether China overestimates its ability to withstand cyberattack and is thus less reluctant to engage in risky behavior.

Designing a SIOP for cyber strikes would be possible, but it would require an intense and continuous reconnaissance effort, extensive target analysis, and the design and testing of special technologies. More important, it may no longer fit with how the United States plans to fight its wars. The SIOP was the ultimate expression of "strategic bombing," the theories and doctrine that grew out of World War I. It assumed that bomber aircraft were unstoppable and could inflict unacceptable damage that would destroy an opponent's capacity and will to resist without having to first overcome the opposing armies in brutal trench warfare.[21] Strategic bombing, before nuclear weapons, was a chimera, as it was both ineffective and counterproductive. Target nations proved to be resilient in repairing infrastructure and industry, and target populations responded to bombing by uniting behind their national governments and increasing their resolve to fight. Resources devoted to strategic bombing came at the expense of supporting offensive operations on the ground. Only nuclear weapons, with their immense destructive capabilities, provided the means to realize the dreams of air power advocates.

The utility of the nuclear SIOP was part of the larger debate over war fighting and deterrence. At one level, the SIOP assumed that the United States would actually use nuclear weapons in a conflict, and that nuclear warfare was not unthinkable. In fact, over the course of a few decades from 1950 to 1970, the superpowers held to an implicit norm against the use of nuclear weapons;

[21] The 1943 Casablanca Directive, which unleashed the strategic bombing campaign against Germany, had as its primary objective "the progressive destruction and dislocation of the German military, industrial, and economic system, and the undermining of the morale of the German people to a point where their capacity for armed resistance is fatally weakened." See also Neville Jones, *The Origins of Strategic Bombing* (London: William Kimber, 1973).

the weapons were in effect “delegitimized” for either tactical or strategic use. National doctrines among those countries with nuclear weapons increasingly reflected implicit agreement that nuclear attack would be reserved for use only in extremis. The weapons retained, however, their important symbolic function as a deterrent and guarantor, leading to the strange Cold War kabuki of planning and refinement of weapons whose use was unlikely.

Cyberattack does not fit the same pattern. Cyber weapons are much less destructive. The flash-and-blast effect of a nuclear weapon would instantly and totally destroy all structures within a radius of several miles around the target. Radioactive debris would continue to pose risk to survivors over a much larger area for days after the attack. Since industrial targets were usually located in or near urban areas, the “collateral damage” could be immense. Even a strike intended to minimize civilian causalities would still kill millions.

In contrast, a cyberattack aimed at crippling industrial infrastructure would most likely disrupt the industrial control systems (ICS) that direct machinery. These systems (also known as SCADA, for supervisory control and data acquisition) are specialized computers with limited functionality. Over time, companies have moved from using dedicated, proprietary networks to communicate with ICS to using the Internet and Internet protocols for their operations. This move has lowered costs and made operations more efficient, but it has also increased vulnerability. In some cases, companies are not even aware of the extent to which their ICS are Internet accessible. The use of wireless connectivity, also attractive on cost and efficiency grounds, increases vulnerability to cyberattack. As societies become more dependent upon computers, the scope and effect of this kind of cyberattack will increase.

Two incidents demonstrate this effect. An accident at the Sayano-Shushenskaya Dam in Russia involved ICS remotely controlled over a computer network. A mistaken command sent over this network caused the turbine controlled by the ICS to explosively self-destruct, killing a number of workers in the immediate vicinity of the turbine and disrupting a critical service. This event was inadvertent, but it could be duplicated in a cyberattack.

The second incident is, of course, Stuxnet. The details of Stuxnet are well known. Stuxnet was a precisely targeted attack combining several advanced techniques and designed to limit collateral damage. There are 4 or 5 advanced “cyber powers” that have Stuxnet-like capabilities, and perhaps another 30 nations are developing them. Stuxnet illustrated the ability of the attackers to circumvent an “air gap” intended to protect the equipment and to cause damage in only a single facility. The lack of collateral damage avoided the negative political reaction from the Iranian population that a bombing raid would have produced. The absence of collateral damage and the reduced likelihood of a negative political reaction make cyberattack different from strategic bombing, where the destructive consequences are visible and immediate.

There is some debate over the ability to predict the scope of collateral damage from a cyberattack, and how uncertainty over collateral damage may restrict use. This has not been carefully analyzed. Attacks like those described above produced little or no collateral damage. The need to design a cyberattack to exploit specific vulnerabilities limits collateral damage. The Stuxnet exploit included malware spread to thousands of systems over the Internet, but this was only a way to deliver for a specialized payload that damaged only the specific target in Iran. Disabling a machine to turn off a crucial service using cyber techniques could produce a collateral effect, but this will be no different than would be the case if you blew the machine up. Certain kinds of cyberattacks pose a greater risk of collateral damage, particularly if the target is using a service shared with other unrelated users, but this risk is not true for all kinds of cyberattacks.

The greater risk of collateral damage comes from a "broadcast" style attack, where the malware is delivered using techniques that infect as many computers as possible. It is not unusual for well-known examples of malware, like "Confickr," to have infected tens of millions of computers. Attacks that disrupt the provision of network services, such as a satellite or public network on which many disparate customers depend, also pose a risk of collateral damage. A commercial communications satellite, for example, services multiple customers, and disrupting service to one may inadvertently disrupt service to innocent parties. Inadvertent damage to third-party networks increases the political risk of cyberattack and could expand any conflict.

This has implications for both "weapons design" (writing software intended to cause damage) and tactics. The risk of collateral damage is minimized if the malware is written only to damage a specific target or if an additional step is required after infection to trigger disruption. Stuxnet, for example, infected hundreds of systems around the world but only damaged one, in Iran. Blanket "attacks" like Confickr or I Love You virus, while they had a global scope and infected thousand of machines, did not really do significant damage of the kind needed for military advantage. Limited commercial disruption does not pose a risk to national security. The intent of these viruses is as a demonstration that goes awry and ends up consuming computer resources or (in other cases) to create remote control capabilities (for bot-nets) and to extract information. Additional code is needed to turn them into weapons. It might be possible to design a virus that would simultaneously disrupt or erase data on thousands of computers—this sort of attack would not be consistent with international law, but that may not deter some classes of attackers.

Cyberattacks may in fact be "single use." They are designed to exploit a particular configuration and set of vulnerabilities. Once the attack has been used, defenders install patches and close vulnerabilities. This limits both the risk of collateral damage and the useful life of the exploit. It also creates incentives for an attacker to strike as many targets as possible in a first strike since the opportunity for follow-on use will be limited. The greatest benefit of cyberattack could come in the opening phase of conflict and then decline in value, unless an attacker had developed many different and independent attack techniques. The nature of cyberattack creates incentives for early use against many targets, a source of instability that will be difficult to address.

A final set of attacks would not produce physical damage but would seek to disrupt services that rely on computer networks. The status of this form of attack is somewhat ambiguous under international law. It is clearly a crime, but not necessarily a use of force that would trigger the right to self-defense. For example, in the December 2008 penetration of the Department of Defense's classified network, the attacker implanted malware that the department was unable to remove for several days. In this instance, the only effect was to penetrate the network; but other malware could have been triggered to erase or scramble stored data. Had this been triggered during a conflict, it would have seriously (and perhaps fatally) hampered U.S. combat operations.

An attacker could look to disrupt and disable services that depend on computer networks, such as automatic teller machines or gas pumps, or erase and scramble financial data or other records. These disruptions are less damaging than destroying critical infrastructure, but an attacker might expect that they could yield military advantage by creating public disorder or political dissent. The degree of disorder may depend on the availability of substitute resources—if one bank's automated teller machines are down, but people can still access other systems, the effect may be minimal.[22]

[22] A 2003 Bank of America automated teller machine outage lasted 24 hours, but produced little reaction from customers. Bruce Schneier, "Blaster and the Greater Blackout," *Salon*, December 16, 2003,

These sorts of disruptions may be more appealing to nonstate actors, whose goal is to make a political point, than they would be to military forces.

It could be possible to attack Internet infrastructure. Global degradation of Internet services by damaging routing services and domain name systems would slow service and introduce a higher rate of error. An attacker could attempt to localize the effect, but overall, the military benefits of such an attack are very limited, given the resiliency of the system and the "payoff" in terms of damage to an opponent's capabilities.

An attacker that sought to create temporary chaos could disrupt large, interconnected national networks for finance, electrical power, gas and oil pipelines, and telecommunications. These targets are attractive, but a single successful attack can affect a broad geographic areas. Or an attacker could focus on specific cities, like Washington or New York, to attempt to disrupt water supplies and government services (e.g., traffic lights), along with the networked targets. Some scenarios for these attacks see them as an effort to distract political leadership or to extend the damage from a kinetic attack.

An integrated plan for a cyberattack could combine all these attacks, prioritize them, and link them to specific targets and timelines. The effect could be highly disruptive and, if used in an all-out surge, even paralyzing for the civilian sector. Nations capable of planning and building such attack capabilities would need to determine when there would be military advantage from such an attack. The immediate question is whether developing a cyberattack SIOP could have a deterrent effect. This assumes that the knowledge that a country possessed advanced cyberattack capabilities would the increase the deterrent effect of existing kinetic capabilities. If a target country already faces conventional forces and nuclear weapons, the gain to deterrence from adding the threat of cyberattack might be minimal.

This sort of integrated, large-scale attack would be both expensive and risky. It would require an extensive reconnaissance effort, simultaneous efforts to develop software to exploit vulnerabilities, and perhaps an unsustainable number of accompanying clandestine activities involving human agents. Skimping on any one of these elements could greatly reduce the effect. That said, the inadequate state of U.S. cyber defenses meant that with some months of preparation, an advanced opponent could reasonably expect to cause widespread, if short-lived, disruption. In addition to disrupting critical domestic services and military command and logistics operations, the goal might be to overload the U.S. political leadership by creating simultaneous "crises" such as blackouts in major cities, disruption of the financial system, or interruptions in the supply of key products like gasoline. One possible scenario for cyberattack would be a surprise first strike launched against targets in the United States, to disrupt the U.S. response to a regional military crisis.

The trade-off a potential attacker would need to consider would be the benefits of short-term disruption compared with the risk of conflict escalation and prolonged post conflict tensions. Opponents who expect their long-term relations with the United States to be hostile in any case may be more willing to accept the risk of prolonged tension. Opponents that assume an escalated U.S. response to their military actions may be less constrained by the risk of escalation. Depending on their connection and status in the "international system," an opponent may be less concerned about the opprobrium that might accrue form this kind of attack—China might find this a greater constraint than North Korea, for example.

Opponents could threaten to use broad cyberattacks against civilian targets to improve their position (e.g., in conflict termination), but a number of factors will condition decisions on this.

http://www.salon.com/2003/12/16/blaster_security/.

A threat to use cyberattack would degrade its effectiveness by warning the target. It could create uncertainty for the opponent, and an attacker might believe that this could lead to improvement in its military or bargaining position, but similar threats to use weapons of mass destruction to constrain U.S. forces, while worrisome, were ineffective. A defender may assess its ability to survive a cyberattack and continue to press military operations as sufficient to allow it to accept the risk, making any threat of limited utility.

Each decision forces consideration by military and political leaders of the balance between the risk of escalation and the military benefit of disruption. If there were perfect knowledge and strong decision-making processes, escalatory attacks would likely be avoided. The emotional pressures of warfare suggest, however, that in the first major conflict were we see the use of cyberattacks, leaders will be inclined to use them more freely against military targets. Assessments of escalatory risk will vary depending on the context and the course of a conflict, and a strike on civilian targets in the opponent homeland, which might seem to hold an unacceptable risk of escalation at the onset of conflict, could seem reasonable in a later phase.

An attacker would need to consider in its planning whether to limit cyberattacks to military systems or to include civilian, nongovernmental targets. A related decision would be whether to limit strikes to in-theater targets, to add supporting forces not in theater, or to expand to general targets in the target homeland. For example, in a conflict over Taiwan, this would mean deciding whether to limit cyberattacks to U.S. forces in the region (including command assets in Hawaii); whether to add supporting elements in San Diego, Seattle, and the Pentagon; or whether to strike nongovernmental targets like the financial system or electrical grid. An attacker would need to decide whether to limit attacks to the disruption of data and computer-dependent services, or whether it would attempt to cause the physical destruction of equipment, gauging their effect against the need to slow or weaken any U.S. response.

The military benefits of an integrated attack might be limited but, as with strategic bombing, this handicap might not be apparent until after a cyberattack has occurred and we can assess its effect. However, as was the case in strategic bombing, where attacker initially assumed that a few raids would lead to panic, rioting, and chaos, and perhaps even to the collapse of the opposing government, a potential cyber attacker might overestimate the disruptive effect. A full-blown, unrestrained, completely successful cyberattack against critical infrastructure and networks might be able to disrupt crucial services and curtail economic activity. Multiple simultaneous cyberattacks would still not guarantee victory but could risk being seen by the United States as an existential threat that would justify a harsh response. There are many examples of militaries attacking targets that were irrelevant to success, but that only inflamed the opponent, so we cannot rule out this kind of cyberattack (which could be very appealing to terrorist groups, should they ever acquire the ability to launch them).

We should ask whether and how cyberattack poses a "strategic threat" to the United States. In its usual context, strategic has come to mean nuclear weapons, posing an existential threat, or a threat to the political independence or territorial integrity of the United States. It is hard to see any cyberattack posing this kind of risk. Perhaps in the future, as computer networks are woven even more deeply into the fabric of our societies, this will change, but the damage a cyberattack could cause is far from existential. That poor cyber security in the face of constant hostile action and potentially damaging attacks does not create an existential threat in the near term does not mean that it is an acceptable situation for the United States. This is an insecure and increasingly unstable environment, where the United States needs new policies to better manage risk.

The Declining Utility of Deterrence

Deterrence is not one of these new policies. Deterrence, a core element of the Cold War Grand Strategy, appeared to have worked so well in that conflict that its advocates are baffled at its lack of utility in current circumstances. The best evidence of the weakness of deterrence in cyberspace comes from the United States, which has some of the most advanced offensive cyber capabilities in the world, yet obtains no deterrent effect from them. Nuclear weapons deterred a potential aggressor. Cyber weapons do not.

The concept of cyber deterrence seems to be another of those instances where the word "cyber" has been appended to a term without sufficient analysis. A careful opponent might even assess the risk of damage from a cyberattack as sufficiently low that if the U.S. response to an attack was limited to cyber means, there might be little or risk and hence no deterrent effect. Cyberattacks are deterred because opponents fear the U.S. military response - what we call general deterrence - not some keyboard-versus-keyboard exchange. Opponents with greater tolerance for risk (or a lesser ability to calculate risk) may not be deterrable at all.

Espionage and crime are not deterrable because we cannot make credible threats. Nondestructive cyber exploits involving cybercrime and cyber espionage, since they fall below the thresholds found in international law to justify the use of force in self-defense, make deterrence largely irrelevant. If we accept that during the Cold War, U.S. nuclear forces had some deterrent effect on Soviet military actions, they were ineffective in deterring espionage, political action, or the use of proxy forces. Deterrence may have channeled Soviet activities into lower-risk actions, but just as Cold War deterrence did not prevent espionage or the use of proxies, deterrence of malicious cyber actions is similarly limited and has no effect on cyber espionage or cybercrime.

Destructive cyberattacks are deterred by general military capabilities; adding a cyber component may increase this marginally. Having a cyber SIOP and communicating this to potential opponents might marginally increase the cyber component of deterrence, but would have no effect on other malicious cyber activity unless there was some way to make it clear that the attack would be a consequence of nondestructive malicious activities. However, threatening SIOP style attacks against other, non-destructive cyber activities would be a significant departure from international law and a troubling precedent for the U.S.

Deterrence rests on a series of assumptions about how potential opponents recognize, interpret, and react to threats of retaliation. One fundamental assumption is that a correct interpretation by opponents of the deterring nation's strategies, capabilities, and responses will lead it to reject certain offensive actions as too risky or expensive. Another fundamental assumption is that the opponent is implacably hostile and only waiting for the correct moment to strike. One dilemma is that if the opponent is not waiting to strike, statements designed to deter it by highlighting the capabilities for a military response (intended to affect its calculation of risk) are likely to be interpreted as aggressive rather than deterrent. A third assumption is that the opponent is risk averse—and politically or religiously motivated opponents are much less likely than government leaders to be deterred by the threat of retaliatory attack.

Deterrence works best in defined circumstances: an opponent with valuable targets that can be threatened with immense damage, an engagement process that allows signals to be sent and risk to be calculated, and some process that allows assessment of the balance of deterrent forces and verification of that deterrence is working. None of these conditions exists for cybersecurity. What could be termed “general” deterrence, the possession of military force sufficient to give pause to an opponent that considers encroaching on vital interests, retains its utility. In contrast, the ability to replicate nuclear-versus-nuclear deterrence in cyber-versus-cyber scenarios does not make sense.

New Classes of Opponents

Deterrence rests on a series of assumptions about how potential opponents recognize, interpret, and react to threats. The fundamental assumption is that a correct interpretation of a threat by opponents will lead them to reject certain courses of action as too risky or too expensive. The problem is that potential opponents may misinterpret deterrent threats while others may be not feel threatened, and are therefore harder to deter.

New classes of opponents complicate the ability to deter by threatening a military response. The United States has gone from facing a single adversary that would often mirror U.S. actions to a collection of near-peer opponents, regional states, and nonstate actors. We cannot safely assume that each set of actors has the same tolerance for risk or will have the same reaction to a deterrent threat. Some opponents may overestimate their strength, and believe that they can succeed in resisting or intimidating the United States. This is an error many have made in the past, and it will likely be repeated again in the future. Some opponents may hold eschatological beliefs about the day of judgment that reduce the effectiveness of threats of retaliation. Politically or religiously motivated opponents are less likely to be deterred by the threat of retaliatory attack. They are inured to threats, and may discount or ignore threats against their forces, cities, or populations.

Other opponents could see a deterrent threat as a challenge calling for an aggressive response. These new opponents already perceive the United States as hostile and threatening, so it may be difficult to increase the level of threat to deter them without heightening tensions. Deterrent threats could actually increase the chance of conflict. With Iran or North Korea, America may find itself in a situation where overly overt deterrence could increase tensions or escalate the risk of conflict rather than deter attack.

Nonstate actors have no capital city or infrastructure to threaten, and their willingness to accept risk is already greater than most nation-states. Nonstate actors do not face the same political constraints that apply to state actions in cyberspace. Some potential opponents may even welcome retaliation, as it could provide justification and expand support for their cause. Against jihadis and other insurgents, the threat to use force will not deter them from attacking. At best, a threat will shape their planning. These individuals have already accepted a high degree of risk in pursuit of their aims and, in the case of jihadis, they believe their populations are already under attack, they lack tangible assets that can be held hostage, and they may not fear death as much.

Even near-peer opponents may not be deterred by threats of retaliation. When Mao Zedong downplayed the potential for American nuclear attacks, this reflected in part a willingness to accept levels of damage that the United States would consider unthinkable—Mao was not “deterred” from pursuing various economic and political policies even thought they cost the lives of millions of people. Mao may have been wrong to consider the United States a paper tiger, but he may have discounted the ability of U.S. forces to inflict unacceptable levels of damage. While China today is a near peer, and likely to be more risk-averse than nonstate actors, China under Mao was close to the regional state opponents we face today, such as North Korea and Iran, which may be willing to accept higher levels of damage in any conflict.

Even a risk-averse China will have a different conceptual framework for conflict and international relations than was the case with the Soviets. China lacks the experience of the Cold War to guide its interpretation of American actions and signals intended to deter. More importantly, some deterrent threats may trigger an escalation of tensions or a violent response. Chinese views are shaped in part by their response to the Century of Humiliation, which leads them to be both more suspicious and more nationalistic. The primary objective of the Chinese government is regime survival, and Chinese leaders see the collapse of the Soviet Union as a cautionary tale about engaging the United States on its own terms. A threat against the political survival of the Communist Party could force China to take extreme actions relatively quickly. America has not adequately assessed the distance between threats that an opponent finds to lack credibility and a threat that an opponent may see as raising existential risk. This distance could be smaller than was the case in the Cold War.

Credible Threats

To be effective as a deterrent, a cyber threat would have to impose an "unacceptable loss' on the opponent. Initial estimates by Robert McNamara calculated that the "unacceptable loss" required for the nuclear deterrence included half of the Soviet Union's industrial capacity, at least two-thirds of its military forces, and perhaps a quarter of its civilian population.[23] This is far beyond the capability of any cyberattack. Nuclear deterrence created a degree of restraint, if not stability, as opponents were afraid of the potentially existential consequences of direct military action against the other. But an attack against a network does not justify a nuclear response, certainly not against another nuclear power and most likely not against a non-nuclear power given the stigma attached to nuclear weapons use. One can ask whether anything less than an existential threat can have a deterrent effect. The lack of destructive capacity and the absence of existential or serious harm undercuts the notion of cyber deterrence.

A credible threat may require the United States to threaten to retaliate in some other domain, but this brings a risk of escalation of conflict. The United States can increase opponents' uncertainty by emphasizing "cross-domain" deterrence, where the response to a cyberattack may not be limited to cyber techniques. The general military capability of the United States, including the cyber component of these general capabilities, is what deters the use of force against it. The goal is to influence the calculations of potential opponents on the risk of taking a hostile action against the United States versus the potential benefits. To change this calculus, the United States could either persuade a potential opponent that there was an increased likelihood of penalty—perhaps by creating international norms that would bring opprobrium upon an attacker or justify some retaliatory action, or it could decrease the likelihood of reward by improving its defenses.

Military force is of limited utility in deterring actual cyber threats. A U.S. military response to espionage or crime would be a strange departure from international norms regarding the use of force. A retaliatory cyberattack (where the intention is to damage or to destroy, rather than exploit), or retaliation using a kinetic weapon for a cyberattack against countries that have not used force against America or against individuals with criminal rather than political aims, could easily be interpreted as an aggressive and unwarranted act by the international community. The result is to cast doubt on the credibility of a retaliatory threat, weakening any deterrent effect.

[23] Robert McNamara, "Mutual Deterrence" speech, San Francisco, Calif., September 18, 1967, http://www.atomicarchive.com/Docs/Deterrence/Deterrence.shtml.

The central flaw in cyber deterrence is that few opponents find it credible that the United States will use military force against espionage or crime, particularly when these actions are sponsored by another sovereign state. A military response to espionage would be unprecedented in international affairs. We can effectively deter the use of force against the United States, but little else. In this, as in anything else, the use of force would involve actions threatening "the territorial integrity, political independence or sovereignty" of the United States. It could involve the use of force or threat to use force (e.g., coercion). The use of force involves physical harm—damage or destruction. Espionage and cybercrime do not produce physical harm and do not qualify as the use of force. This makes threats to use military force against them not credible, leaving the United States in the position of having one of the most powerful cyberattack capabilities in the world and finding it of little use in deterring cyber exploits (how it could be used for defense is another matter).

One suggestion that has been made to fix the lack of credibility is to lower the threshold for the use of force for cyberspace, so that we could consider espionage or cybercrime as actions that justify a military response. This would be exceptionally shortsighted. The United States itself has an immense espionage effort, including the use of cyber espionage. We do not wish to constrain our own ability, nor do we wish to justify other nations using military force against us. Deterrence is less effective in an environment where the United States has more to lose than its opponents.

Another suggestion has been that the United States adopt new procedures to justify the use of force. When we discovered an exfiltration program or a computer penetration, we would first warn the government with jurisdiction that we were going to take action against the offending computer located in their country if they did not themselves take action against it themselves. If no action were taken, the United States would be justified in reprisal. Under this approach, if we found a computer in China responsible for espionage, we would go to the Chinese government and tell them that if they did not take action we would damage or disable the computer.

There would be benefits from confronting a nation we believe is tolerating if not directing cyber espionage (although the risk to future collection from disclosing persuasive evidence would have to be considered). The problem with this "warn before shooting" idea is what if the other government says no? Nations are far more protective of their sovereignty than is generally recognized, and warning a nation that you intend to violate its sovereignty in a way you consider to be justified may not produce the expected results. The international community is very slow to sanction the use of force against a sovereign state, even in much more compelling circumstances (e.g., Libya or Somalia).

One way to make this linkage would be to warn an opponent that if malicious cyber actions continued unchecked, this would trigger a cyberattack in response. The dilemma would be in preventing this from escalating to a broader military conflict using kinetic weapons. An astute opponent would imply that a cyberattack in response to espionage would trigger a kinetic conflict, to complicate planning and deter use. Threatening a large-scale SIOP-style attack in response to espionage would likely not be perceived as credible. Even precise, targeted attacks that damaged or destroyed a single server in another country could, however, be regarded as the use of force and as an illegal act. To put this in a physical context, if pirate vessels operate out of Hainan Island, the United States could not assert the right to sink them in harbor. A long series of diplomatic steps would be required first and, absent Chinese cooperation, would lead ultimately to an ultimatum—stop the pirates, or we will use force. This is a very-high-stakes game that would certainly escalate tensions and could easily trigger unexpected retaliation.

Reprisal is not the same as creating consequences for malicious cyber actions. We should note that all state parties engage in espionage and that each feels itself to be aggrieved. Misinterpretation is likely to increase if one party engages in reprisal and the other does not realize that the act is intended as a reprisal. Reprisal depends on the attacker feeling that it is being punished and should stop. It is just as likely that an attacker may not be dissuaded by a single reprisal or could see it as the onset of an exchange of attacks. In these circumstances, reprisal is more like to escalate conflict than to end it.

It is also not clear that a United States–initiated cyber exchange, even assuming that it could manage the risks of escalation, would play out in its favor. We cannot assume that after having caught an opponent, warned them, and then retaliated against them, that they would see this as justified or that they would not respond in kind. There may be an asymmetric vulnerability, with the United States depending more on computer networks than an opponent, and this means that the United States would come out worse in any cyber exchange. Even if this asymmetry does not exist, an opponent may believe the United States has greater vulnerabilities. Tit-for-tat cyber exchanges are not conducive to stability.

One unique aspect for reducing the risk of cyberattack, or any malicious cyber activity, is the opponent's calculus not of the risk of retaliation but of the risks of discovery. Since attribution is not perfect, an attacker may calculate that they can act safely. If an attacker calculates that the United States will be unable to identify them as the perpetrator, they will perceive less risk in carrying out an attack. This was not an issue for nuclear deterrence, where both missile trajectory and radiological evidence eliminated the attribution problem. Increasing the attacker's uncertainty about whether or not they are anonymous will deter attacks. To help opponents better calculate risk, and to increase its ability to "deter" cyberattacks, the United States could make explicit statements on its own attribution capabilities. Announcing that you have knowledge of a potential opponent's plans has a similar effect, although it runs the risk of alerting them to your ability to collect against them.

Uncertainty

Deterrence strategies in the Cold War accepted a large degree of collateral damage as necessary for threatening nuclear retaliation. Nuclear strikes would have harmed civilian populations in both NATO and Soviet Bloc countries. But these strikes were reserved for extreme situations when sovereignty had been clearly violated by military force. The extent of collateral damage for nuclear weapons was in some ways easier to predict than is the case in cyber conflict—the blast and radiation effect of nuclear weapons is limited to an area around impact; in cyberspace, collateral damage may not be contiguous with the target or even located in the target country.

In the Cold War, there was clear attribution in that allowed for both credible threats and for "signaling" and tacit understandings between opponents on "redlines" and thresholds. America lacks that clarity in cyber conflict. More important, an anonymous attacker may not lose anything at all since its identity is unknown and retaliation is impossible. Weak attribution makes traditional deterrent concepts—those based on the threat of reprisal for an attack (either counterforce or counter value)—largely irrelevant in cyberspace.

Uncertainties over attribution and collateral damage will limit the willingness of political leaders to authorize cyber counterstrikes and erode the ability to make a credible threat. Until we can predict with confidence the scope of collateral damage, a cyber counterstrike could easily damage an ally or neutral party. The interconnectivity of cyberspace makes predicting collateral damage difficult. Uncertainty about the scope of collateral damage involves both unintended effects on the target and also possible damage to third-party networks connected to or dependent upon the target network. Disabling or disrupting one network may affect third parties; for example, an attack on an opponent's network might accidentally degrade a neutral nation's satellite or telecommunications services. Uncertainty about collateral damage may hobble deterrence in cyberspace, by reducing the willingness of political leaders to incur the risk of widening a conflict or creating unfavorable political consequences.

Inability to Extend General Deterrence

The United States has successfully deterred attack by conventional military forces. More than this may not be possible. The United States today has overwhelming military force at its disposal. Its opponents strive to avoid conventional warfare or strategic exchanges with the United States, as these would be costly and likely result in defeat. Near-peer opponents, either by intent or by a desire to avoid risk, may shy away from actions that could reasonably justify the use of military force by the United States. While near-peer action in cyberspace is damaging, it has remained below the threshold of an act of war, has not involved violence, and could not be considered an act of aggression under international law. America can claim success in the ability of its extraordinary military capability to deter direct military attack, but in other areas, including cyber conflict, opponents may be unresponsive to threats to use military force.

This may also point to the existence of implicit thresholds for cyber conflict—if attackers limit their actions to espionage, which is generally not regarded as an act of war, there is little chance that the victim will undertake retaliatory attacks. Deterrence by threatening retaliatory attack does not increase security in cyberspace.

Deterrence in cyberspace is limited because we have not adequately assessed what combination of cyber capabilities, defensive measures, and international agreements will make the United States and its allies most secure. It would be useful to undertake a larger strategic calculation, preferably in a public dialogue, to determine the weighting and balance among offensive, defensive, and multilateral efforts in cyberspace that best reduce the risk of cyberattack. The constraints that limit the value of deterrence would change with the onset of conflict. Experience suggests that even without extensive tutoring by the United States to develop a shared conceptual framework and lexicon, there is some fundamental level of deterrence that new classes of opponents will intuitively understand. Deterrent threats, especially cross-domain deterrent threats that are credibly communicated, could be useful in shaping opponents' choices on tactics, targets, and weapons in conflict.

Moving Past Deterrence

The appeal of deterrence is unrelated to its effectiveness. In addition to nostalgia for the simplicities of the Cold War, deterrence has politically attractive attributes. Deterrence is a unilateral action—it does not require cooperation with another nation. Deterrence does not require engagement—other nations deduce your intent from indirect actions, signals, and other behavior. Deterrence relies on offensive capabilities and avoids the many messy issues involved in creating an effective cyber defense. While deterrence may be politically attractive, it is severely limited in its benefits for cybersecurity.

Deterrence cannot be the centerpiece of any U.S. cybersecurity strategy. Numerous problems reduce its utility. If the attributes of classic deterrence were symmetric vulnerabilities, highly destructive weapons, and existential threats from a peer opponent, the United States could believe it could manage stability by ensuring a rough equivalence of forces so that the Soviets never perceived a moment when the benefits of attack outweighed the costs. This may not have actually been how the Soviets made their strategic calculations, but it provided a useful framework for U.S. planning and strategy. The same is not true for cyberattack. Vulnerabilities are not symmetric (particularly against opponents other then China or Russia), and cyber weapons do not pose an existential threat similar to that posed by nuclear weapons.

Deterrence is politically attractive. It does not require engagement with other nations and winning their consent to norms or constraints. It does not require the investment and regulation required for better defense. But its utility in cyber conflict will be much more limited than in the past. Deterrence is itself inadequate and must be buttressed by political actions that go beyond classic, force-based deterrence. This points to the need for engagement, norms, and understandings to create a common framework for nations to think about cyber conflict.

Given the difficulties of making credible threats, reducing the likelihood of cyber attack will require something other than increasing military capabilities. The likelihood of attack could be reduced through stigmatization—the creation of a credible international norm that says some forms of attack run counter to accepted international behavior. The use of nuclear weapons was stigmatized, but it may be more difficult to stigmatize less destructive forms of attack or to extend stigmatization to all forms of attack in a particular domain.

Cyber defense would be reinforced by multilateral understandings on acceptable behavior in cyberspace—explicit norms or obligations. A norm that establishes a state's responsibility for the private actions of its citizens could make it more difficult for Russia to plausibly deny its involvement in attacks on Estonia. Just as nations feel a degree of constraint from norms and agreements on nonproliferation, establishing explicit international norms for behavior in cyberspace would affect political decisions on the potential risks and costs of cyberattack.

The United States could increase the likelihood of success for a defensive strategy by indicating constraints on its own military activities, as it did with the Soviets during the Cold War. Fear of U.S. capabilities drives some opponents' actions, and misperceptions of U.S. activities are used by opponents in internal debates to justify a response to U.S. aggressiveness or "hegemony. The United States would need to identify not only what it would be willing to give up but also what it would want from opponents in exchange. This kind of negotiating process would help to establish bounds for cyber conflict, making cyber conflict easier to prevent or manage.

The most useful element of deterrence might be the ability to "signal" an opponent about potentially risky behavior. These signals could include implicit warnings created by changes in force status or readiness posture concern over opponents' behavior, by developing tacit understandings on "redlines" and thresholds, by implicit or explicit understandings among potential opponents, and by public statements about intentions. Signaling cannot occur in a vacuum, however—the opponent must sufficiently understand U.S. doctrine and practices so that it can correctly interpret the signal. Common understandings are necessary as they allow for both credible threats and for tacit communication with an opponent. Developing such understandings could make cyber conflict easier to prevent or manage.

Signaling in Cyberspace

Signaling, like deterrence, is another inheritance from the Cold War that people now try to apply to cybersecurity. Signaling is tacit and indirect communication. The United States decides to take an action that does not directly engage the opponent but that signals its intent or concern. The opponent must understand that this signal made in response to its action. An opponent's ability to correctly interpret a signal depends on experiences and an accretion of authoritative public documents on doctrine and policy that reinforces any signal. These are lacking for cybersecurity.

Both signaling and deterrence share a common problem. They are reactive and unilateral techniques derived from a bilateral confrontation involving nuclear weapons that posed an existential threat to the survival of the United States. Engagement with the Soviet Union was difficult and initially impossible—as George Kennan said in 1947, the Soviets were amenable only to the "logic of force." Soviet society was isolated, and contact between the West and the Soviets was infrequent. Signaling avoided explicit statements that could be interpreted as threats, allowing nations to exercise caution when dealing with the risk of nuclear war, and provided an indirect means of communication when direct contact was not possible.

The classic example of Cold War signaling has a Soviet missile submarine move closer to the United States (this meant a shorter flight time for a missile and less warning time, which reduced stability by increasing the chance of a surprise attack). In response, the United States might visibly move bombers to a higher readiness state. Soviet reconnaissance satellites would detect this change in status, and the submarine would draw away from the coast.

This kind of signaling will be difficult in cybersecurity. What would moving to a higher state of alert entail? America is not in a bilateral conflict; nor does it face an existential threat. Engagement on many levels with potential opponents other than Iran and North Korea is routine. This does not mean that the United States can ignore signaling, as foreign audiences will interpret its actions as signals of intent, but it means that explicit, formal communication on concerns should be preferred.

Signaling raises the problem of misinterpretation. Even overt, explicit messages can be misinterpreted, and the likelihood is greater when using actions to implicitly signal. Foreign audiences come at communications with different beliefs and expectations. Too subtle a signal may be missed entirely. Or a signal intended for one situation may be interpreted by the audience as applying to something completely different.[24] This is a second-best tool for those audiences with which direct communication is possible.

A review of documents from Soviet archives made available after the Cold War shows that the deterrent message the United States thought it was sending was often not the message the Soviets received. The possibility of miscommunication exists today. Potential opponents may misinterpret signals as expressions of hostile intent, or they may discount them. The risk of misinterpretation is high. Beliefs about America's true intentions and strategies will shape how other nations interpret messages and signals. For example, Chinese officials have said that the United States has a grand strategy whose goal is to preserve U.S. economic, military, and technological hegemony. This means that actions the United States believes to be benign may be interpreted as malicious. In either case, what is intended to be a signal may in fact be destabilizing and increase the risk of conflict.

[24] Scott Douglas Sagan and Jeremi Suri, "The Madman Nuclear Alert: Secrecy, Signaling, and Safety in October 1969," *International Security* (Spring 2003), http://live.belfercenter.org/files/sagan_and_suri_spring_2003.pdf.

A good example is found in a statement by a Chinese official, who said that the U.S. Department of Homeland Security exercise Cyber Storm, where the United States sought to improve a coordinated response to cyberattacks, was like “missile defense.”[25] Asked to amplify this, the Chinese official said that the intent of missile defense was to cancel out China’s strategic deterrent, and that the true intent of Cyber Storm was to give the United States impunity in cyberspace so it could strike China with cyberattacks without fear of retaliation.

The reaction to the release of the U.S. Department of Defense’s 2011 cyber strategy illustrates the problem. The strategy echoed existing “declaratory policy” by saying that the United States would consider a range of responses to a cyberattack, and that if it determined that the cyberattack was the equivalent of an act of war, this could include the use of military force. An earlier statement by a senior military official had already indicated that the United States was not limited to cyber means in a response to a cyberattack—his colorful phrases was “If you shut down our power grid, maybe we will put a missile down one of your smokestacks.”[26] While the idea of cross-domain deterrence is sound, since it is in fact a reference to general deterrence and warns opponents that the United States is not limited to a relatively feeble cyber response, the strategy and the statement led to widespread misinterpretation. Some of this misinterpretation may have been intentional by those hostile to the United States, and misinterpretation was reinforced by the unfortunate tendency of the media and the public to call every malicious action in cyberspace an attack.[27]

It is useful to ask how an opponent would interpret U.S. actions as a signal about its strategy and intent. If an opponent believes that the U.S. intent is to contain and undercut its regime and national power (and leaders in both Russia and China share this suspicion), opaque or ambiguous statements are unhelpful. The chance of misperception is increased because democracy promotion threatens both regimes. This unavoidable tension is not unmanageable—ways were found to reduce the risk of military clashes during the Cold War while continuing to promote democracy, nor should the U.S. abandon democracy promotion, but recognition of this factor should shape U.S. expectations for the likelihood of agreement and the scope of misinterpretation by primary opponents.

Signaling is appropriate when the subject is deemed too sensitive for direct communications, or when the United States does not wish to elevate an issue and perhaps increase tensions. This is a normal part of diplomacy and provides a useful emollient in international relations. However, a decision to use signaling would require careful calculation as to what actions would be seen a proportional to avoid increasing the risk of conflict.

For example, the United States could adopt a more aggressive posture in the South China Sea to signal its displeasure with Chinese cyber espionage. Even if this was carefully explained to the Chinese, it risks becoming destabilizing as they may see it as a disproportional response that threatens its territory and sovereignty. When the Chinese complain about U.S. air and naval intelligence gathering off their coasts, it is useful to remind them that espionage is not limited to aircraft and ships, but the chances of them misinterpreting a more aggressive posture by the U.S., given their beliefs about U.S. intent, is high. Signaling requires preparation (through public statements and direct communications that guide the opponent), and careful calculation of what kind of “signal” would not increase tensions.

[25] Comments made at Track II Discussions, December 2010.

[26] Siobhan Gorman and Julian E. Barnes, “Cyber Combat: Act of War,” *Wall Street Journal*, May 30, 2011, http://online.wsj.com/article/SB10001424052702304563104576355623135782718.html.

[27] Katrina Timlin, “International Coverage of the Department of Defense’s Cyber Strategy,” CSIS Blog, July 18, 2011, http://csis.org/blog/international-coverage-department-defenses-cyber-strategy.

If we accept that behavior is a signal to other nations, the United States has already emitted a plethora of signals, perhaps inadvertently. The shaping event may have been a series of statements in the early 2000s by U.S. military forces that America would seek to dominate cyberspace and deny its use to its opponents. In the context of the invasion of Iraq and similar statements about space strategy, many nations concluded that the greatest risk to stability in cyberspace came from the United States. These statements built upon foreign perceptions of pervasive U.S. global intelligence collection. In the Cold War, there was a degree of tolerance for the United States' actions (except from the European left) because it was seen as defending democracy against a totalitarian regime, an inheritance in some ways of the U.S. role in the anti-totalitarian alliance in World War II. This tolerance has been weakened. Although the 2009 change in administrations greatly reduced fears about U.S. intentions in cyberspace, some audiences remain skeptical or suspicious.[28]

The kind of specific linkages used in Cold War signaling may not be that useful for cybersecurity. The United States would benefit more by making a series of explicit statements on policy, strategy, and potential responses that are not linked to any specific event or aimed at a specific audience. In order to do this in a credible fashion, however, the United States would need to identify what responses to cyber actions it is willing to actually undertake, and it would need to clarify its current strategies.

The Case for Explicit Redlines

One fundamental question for U.S. policy is whether the United States increases the risk of malicious cyber action if it announces explicit thresholds or redlines. The fear is that defining a threshold will tell opponents exactly what they can get away with. Since U.S. opponents are already getting away with so much, it is not clear that U.S. interests would be much worse off in this case. The counterbalancing risk is that the United States increases the chances of an opponent's miscalculation by not setting out clear guidelines for their behavior.

The experience of deterrence and signaling shows that the chance of misperception and misinterpretation by a potential opponent is much higher than expected. What is said is not necessarily what is heard or what is believed. States probably underestimate the level of misperception and overestimate the degree of understanding of their concerns by potential opponents. Finding ways to increase clarity and precision to opponents' thinking on cyberattack would reduce risk, and this should be a goal for U.S. policy.

There is already an implicit threshold regarding cyberattack. Despite countless penetrations of U.S. networks, there has never been any "attack" that the U.S could consider, under international law, as the use of force or an immediate threat to U.S. territorial integrity, sovereignty, or political independence. It is unclear if this is an intentional decision by opponents, but they probably expect that some actions (espionage and crime) as unlikely to provoke a damaging U.S. response, whereas other actions that inflict physical harm bear much greater risk.

The implicit threshold among states is to avoid cyber actions against each other that could be interpreted as the use of force if that is not their intent. Efforts to "lower" the threshold to discourage other malicious cyber actions do too much violence to existing international norms and law. This does not mean that the United States is without recourse. The key to doing this is vigorous

[28] Robert Wright, "President Obama's Hypocrisy on Cyberattacks," *The Atlantic*, June 3, 2012, http://www.theatlantic.com/international/archive/2012/06/president-obamas-hypocrisy-on-cyberattacks/258016/.

response consistent with international practice. Espionage does not justify military reprisal, but there are (as discussed in the next chapter) a range of actions that can be taken in response.

For meaningful effect, there must be explicit statements about specific incidents, not vague or general pronouncements. A failure to comment on specific incidents undercut the general statements that the United States has made about its intentions for cybersecurity. Explicit public statements shape foreign perceptions of U.S. cyber capabilities and intentions and indicate how the United States will respond to cyberattacks, how the United States will use cyber militarily, and how it will apply the laws of armed conflict. These statements can set a precedent, as others may copy U.S. doctrine and policy, and they can be stabilizing if interpreted correctly. They can help reduce fear of the United States and shape foreign efforts to build cyber capabilities.

In the NATO context, for example, there would be benefit from generally clarifying for an external audience when Article 5 is triggered by a cyber incident. This threshold setting increases the likelihood that an opponent would calculate that the risk of a cyber exploit was too great. NATO could publicize the principles and criteria it would use to make a decision about how to respond to some future cyber incident against its own networks or the networks of a member state, to establish what behaviors are unacceptable without getting into precise detail. Describing the conditions under which a cyber incident changes from a threatening, politically coercive action (as we saw in Estonia) to being the equivalent of the use of force can manipulate an opponent's planning in beneficial ways.

A clear declaratory policy—public statements such as the one found in the U.S. president's May 2009 speech or in later policy documents—would help reduce risk. America could begin to identify a threshold for when actions in cyberspace could be considered an act of war if it accepted that an action in cyberspace that produces the equivalent effect as physical sabotage begins to rise to that level. It would be an exceptionally serious matter if a nation sent agents or soldiers across another nation's border to blow up a pipeline or power station and a similar action in cyberspace, unlike crime or espionage, could justify a military response:

We reserve the right to use all necessary means—diplomatic, informational, military, and economic—as appropriate and consistent with applicable international law, in order to defend our nation, our allies, our partners and our interests.[29]

The most stabilizing declaratory policies are multilateral, where nations jointly endorse principles and commitments on their intentions and behavior. Examples include the 1963 "Declaration of Legal Principles Governing the Activities of States in the Exploration and Uses of Outer Space" and the "Declaration of Principles Guiding Relations between Participating States" in the Helsinki Accords. The process of identifying and winning agreement to such principles might have to be incremental, beginning with unobjectionable ideas and moving over a period of years to the more difficult, but this process will reduce risk and increase stability.

Opponents also assess and draw conclusions from their experience in undertaking malicious cyber actions against the United States and from the U.S. response. They could easily believe that the United States accepts cyber espionage as part of the "game." The United States engages in it itself, and has not objected to the actions of others when they use cyber techniques. Perhaps this is interpreted as a signal of acceptance or indifference. It is certainly not a signal to stop. There is a delicate issue at the center of this. It is not in the U.S. national interest to refrain from espionage

[29] White House, "International Strategy for Cyberspace," http://www.whitehouse.gov/sites/default/files/rss_viewer/International_Strategy_Cyberspace_Factsheet.pdf.

against military and political targets. It did not make sense to say that "Gentlemen do not read each other's mail" in 1929, and it makes even less sense now.[30] The United States will engage in this kind of espionage, and so will its competitors and potential opponents. The goal is not to end espionage but to modulate and reduce it.

"Strategic ambiguity" is counterproductive for cybersecurity. Opponents base their assessments of U.S. policy on U.S. statements and actions. When ambiguous statements are accompanied by a lack of action, opponents could interpret this as indicating that the United States will tolerate malicious actions. They perceive consistency, not ambiguity, since the lack of response has been a feature of U.S. policy for more than a decade. Ambiguity, strategic or otherwise, increases the risk of miscalculation. The goal should be to significantly reduce ambiguity about intentions, redlines, and responses. Uncertainty as to how the United States could respond tends to be confused with opponents' uncertainty as to whether their attack will succeed in accomplishing its objectives. The best outcome would be one where the opponent knows the cost of an attack precisely but is unclear about the potential benefits.

[30] This was attributed to Secretary of State Henry Stimson when he withdrew funding for the first U.S. cryptanalytic intelligence office.

3 CONSTRAINING CYBER ESPIONAGE

Espionage can change the balance of power among states. In this, the U.S. is at something of a disadvantage because it does not engage in economic espionage, the theft of information from private companies by states or their proxies with the intent of improving competitiveness and speeding the development of technology. Offensive cyber capabilities will not deter or prevent cyber espionage.

Espionage in cyberspace is not new. If we define it as remote penetration of an opponent's computer network to extract data, it predates the Internet, with incidents as early as 1982. Given that one operational goal of espionage is to be undetected, or if detected to confuse any observers, and its general goal is remain secret, we should not be surprised at the general low level of discussion of this topic. The best way to think of this is that the Internet has created a golden age for intelligence collection and, in the event of conflict, for decisive knowledge of an opponent's intentions and capabilities.

Any analysis of cybersecurity needs to accept the fact that cyber espionage will continue. International law offers little recourse—the practice among nations has been to discreetly look away when the topic of espionage arises – it is in no country's interest to see this change. The United States accepts that countries will engage in economic espionage, but the question is what kind of international framework of understandings would manage and restrain these activities so that espionage does not reach intolerable levels. China and other nations see cyber espionage as a strategic tool. Illicit access to U.S. economic and military technology is immensely valuable (how well these nations are able to use the intellectual property they acquire is another matter). If the United States fails to protect its networks and fails to object to espionage, there is no incentive for them to stop. Managing cyber espionage will require the United States to create consequences that do not involve the threat of a military response, because this threat is not credible.

An understanding of signals intelligence is critical for a coherent discussion of cybersecurity. Marconi invented the radio in 1912. Navies immediately saw the potential and adopted it to coordinate their far-flung fleets. Naval operators noticed that they received not only their own signals but also the signals of other fleets, unencrypted and accessible to anyone with a radio receiver. This happened in months. The lessons we can draw from the early days of radio is that intelligence opponents are quick to exploit vulnerabilities and commercial technology is generally inadequate to repel professional intelligence services.

Militaries had quickly learned that anything transmitted over the radio in the clear was a gift to their opponents and sought to encrypt their signals. An encryption device takes plain text and turns it into sequences of letters that are unintelligible to anyone who does not have a "key." Now, to exploit radio signals, an operator needed not only to intercept a message but also to be able to break its code. Encryption is hard to do, and it is hard to tell when it is not working. The United States, for example, was able to break Japanese code during the Washington Naval Conference, providing its negotiators with Japan's internal communications on negotiating strategy. The lesson from this is that those who assume they are safe are in for unpleasant surprises, and that not seeing an opponent's activity does not mean, in this world, that it is not occurring.

A more telling episode occurred during World War II with the penetration of secure German radio communications. In the 1920s, Germany took a commercial encryption machine, Enigma, and adopted it for military use. Polish agents stole one of the machines from Berlin and developed primitive computers (called Bombes, because of their shape) to help break Enigma's codes. With the fall of Poland, the Polish cryptographers moved to France. With the fall of France, they moved to Britain. The British took the Polish advances, developed their own, more advanced computers (known as Colossus), and set about finding the key to German codes—a task that had to be repeated every night because the Germans changed the settings of the Enigma machines every day to create a new encryption pattern. The British named their effort Ultra.

There are several lessons from the Enigma experience. A state can draw on a range of resources—human agents, scientists, ships, aircraft, and now satellites—to achieve intelligence success, and the best opponents blend intelligence capabilities to achieve their goals. The Germans wondered if their signals had been penetrated, but they dismissed these concerns because they were convinced of their superior technology. The advantage the British gained had to be recreated on a daily basis. Signals intelligence is a race between measure and countermeasure. Human error plays a crucial role—German radio operators used the same heading to begin their messages, providing a vital clue to the British—and software, when the mistake is coded into a program, expands such errors a thousand times. That said, even having access to the most secret German plans did not ensure success, given that intelligence is only as good as those who implement it—the best example being the successful German airborne invasion of Crete, where Ultra furnished precise details of the attack but the local commander failed to act on this information. Despite these lapses, reading traffic the Germans believed to be secret provided immense military advantage.

The pattern for cyber espionage has been to first gain access and control to an opponent's network. Phishing is currently a favored technique, where an innocent-looking email that often spoofs the email address and appearance of a friend or associate contains a document or link that, when opened, implants malicious software on the network. Another technique is to infect an external device (a camera, smart phone, or thumb drive). When the user plugs it into a network, the malicious software implants itself. This is an obvious and successful means of beating an air gap or penetrating a classified network. An attacker can look for misconfigured systems that are accessible to the Internet, systems were the operator has forgotten to change the default password, for example. Misconfiguration is surprisingly common.

An attacker can exploit undiscovered flaws in commonly used software, which are also surprisingly common, for access. An attacker can acquire a legitimate password or credential through digital means or through conventional human espionage and impersonate a legitimate user. Complex operations combine several of these techniques. Stuxnet, for example, stole digital credentials from an Asian firm, took advantage of several undiscovered flaws in an operating system, and used an external device to overcome an air gap and implant the malicious program.

The intelligence services of major countries can combine cyber techniques with the actions of human spies and traditional signals intelligence to create successful operations. These complex operations are much more difficult to prevent. As software slowly improves, the difficulty of penetration will increase and amateurs will find it harder to play this game, but advanced services, with their resources and their combined technical means, will retain an advantage. The task of cyber espionage will become more difficult, and a sophisticated opponent will still be able to achieve success.

Once the attacker is in the network, they have several choices. They may sit and siphon off traffic over a long period of time—the Canadian company Nortel had industrial spies sitting on its networks for several years. They may implant a few lines of code in some unobtrusive location as beacons—guide points for later reentry and disruption. Or they could do the equivalent of a smash-and-grab, by breaking in and immediately extracting thousands of documents in a short period (usually overnight, when the chances of detection are less)—this has happed to the Department of State and Department of Defense, and to many companies.

In fact, the primary challenge for sophisticated intelligence agencies is not the collection of data, so porous are Internet-based systems, but the ability to store, process, and analyze the data they have acquired. Most processing and analysis are done by computer, with immensely powerful machines using complex algorithms to sort through traffic and identify the tiny fraction, less than one-thousandth of 1 percent, that needs to be looked at by human analysts. Advanced systems allow analysts to go back and search stored traffic for related messages that may not have been selected for review or to collate information from other sources (human agents, for example, or other technical collection techniques). A targeted collection, where an attacker knows exactly what they are looking for, poses less of a problem for information management.

The theft of military technology has been part of espionage for centuries. State-sponsored economic espionage is a more recent development. The United States is at an asymmetric disadvantage in economic espionage, in which it does not engage. Companies routinely copy each other, but these disputes are left to patent law and the courts. In the 1950s, however, the Soviets discovered that their economy lagged behind the West in its ability to develop new technologies, including commercial technologies. Their solution was to create formal state programs to steal technology that Soviet industry can copy—military technology, commercial technology with military applications, and even commercial technology that creates only economic benefits.

Current Russian cyber espionage efforts seem to focus less on economic espionage and more on traditional military and political concerns, such as probing NATO networks for military plans or engaging in sophisticated information operations. Russian espionage will follow the larger pattern of its international actions, where the constraints on its power and its general decline force it into the role of "spoiler," a possible example being the assistance of Russian services in detecting and countering cyber efforts against Iran like Flame and Stuxnet. Russian cyber espionage will also be used to support efforts to maintain influence in the "near abroad" and to assist Russian organized crime activities in Europe. These are all important but do not pose the same level of strategic challenge as espionage activities by China.

Technological and economic espionage was not a problem for China under Mao, who isolated China from outside contact and believed that the mobilized masses could overcome superior technology, but when more pragmatic leaders took over they found an immense lag that affected China's military and economic power. In the mid-1980s, three leading scientists wrote to Deng Xiaoping stating that if China did not make an effort to remedy this technological lag it would fall irretrievable behind the United States and never achieve its goal of leapfrogging the West in military and economic capacity.

Deng responded in March 1986 by creating an immense series of programs to build a strong science and technology base for China. He also made the illicit acquisition of technology a central element of China's economic opening to the West—it was part of every business negotiation and led to long-running state espionage programs targeting Western firms and research centers. This technological espionage has carried over into cyberspace, as the Chinese discovered that the Internet gave them unparalleled access to poorly secured Western networks.

The Chinese justify these programs as something owed for the Century of Humiliation under French and British imperialism, as a something the world should accept given China's poverty, and given the belief among some Chinese (especially in the military) that the United States was implacably hostile. China is not alone in conducting this kind of economic and technological espionage. Russia, Israel, and France have similar programs, and Japanese companies once faced similar accusations. The scope of China's programs, however, surpasses those of these nations, and the global connectivity provided by the Internet has contributed to this massive success.

Sometimes we hear that the United States should not object to industrial espionage by China as America itself did this in the 19th century to Britain. This is a deeply flawed argument. First, Americans could only copy individual books; Chinese can steal the digital equivalent of millions of books in a few hours. Second, nations, including the United States, agreed that the protection of foreign intellectual property (IP) best served the public interest and developed formal agreements among states to create these protections; these agreements did not exist in the 19th century. Today, unlike the United States of the 19th century, China and other nations are violating their international commitments by tolerating IP theft; and by establishing state programs to steal IP (which the United States never did). Perhaps most important, the United States was a net contributor to the global stock of knowledge during the 19th century, with its inventors creating steamboats, the telegraph, the cotton gin, and many other innovations, which other nations copied freely in the lax IP environment of that time. The current perpetrators of industrial espionage have made no such contributions.

More important, however, economic espionage reflects deep political and cultural issues. Chinese attitudes toward IP ownership, always ambivalent, were not strengthened by 30 years under Mao and communism. The Chinese realize that weak IP protection hurts their ability to create new technologies, but getting compliance from the free-wheeling Chinese economy will be an arduous task. Chinese economic espionage reflects deeper problems with the protection of IP and may be so pervasive that some Chinese analysts fear that it may be uncontrollable. Stealing Western technology compensates for this inability to create, but it also reinforces the very trend that hurts China's own technology efforts. Efforts to mandate the use of Chinese technology, such as the "Indigenous Innovation" program, have largely failed (even if China has not abandoned them. There is an internal conflict in Chinese policy between the long-running national effort to illicitly acquire technology from Western companies and the desire to develop China's own ability to create new technologies.

Cyber espionage damages the long-term technological leadership and economic competitiveness of the United States, particularly since the biggest perpetrator is China, a nation that can veer toward confrontation and approaches economic relations in "zero sum" terms, which means that for China to gain the United States must lose. China's economic espionage, at its worst, could "hollow out" key parts of the American economy (and the economies of other nations). Assessing the credibility of this requires, however, weighing the effect of U.S. domestic policies on its economy (as they may be more damaging than economic espionage) and the period of time required for irreparable harm to occur.

The theft of IP and confidential business information—economic espionage—appears to cost developed economies much more and does pose a threat to security, by undermining the military advantage provided by technology and by damaging economic competitiveness. One way to think about cyber espionage is that Americans bear the cost of car crashes as a trade-off for the convenience of automobiles; similarly, they may bear the cost of cyber espionage as a trade-off for the benefits to business of information technology and from operating in foreign markets. That does not mean it is not in the national interest to try to reduce cyber espionage, and of course the theft of sensitive military technology creates damage whose full cost is not easily quantifiable in monetary terms.

Estimates of the dollar value of annual losses to businesses from cyber espionage show a tremendous range, from a few billion dollars to hundreds of billions. It is difficult, however, to assess the accuracy of these estimates. Estimates are based on anecdotes and extrapolation. Problems with valuation and underreporting complicate the estimation process. Companies conceal their losses, and some are not even aware of what has been taken. This means that any number we develop for the loss from economic espionage is an estimate, extrapolated from related or partial data and based on underlying assumptions.

We do not want to assume that losses are distributed evenly across all sectors of the economy. We know that state-sponsored espionage will focus on areas of concern to governments: military and advanced technologies in aerospace, materials, information technology, and sensors, financial data and energy related information. Semiconductors and solar energy have been prime targets. Private hacking, however, could engage a much broader swath of companies. One interview with intelligence officials told of a small U.S. furniture company being hacked and losing its IP. There is no possible national security benefit to this, and there are similar stories involving breakfast cereal, running shoes, automobiles, and soft drinks. This broad range probably reflects not only commercial interests but also an official policy to encourage the illicit acquisition of technology as a way to promote economic growth.

Extracting information through espionage does not mean there is immediate benefit to the acquirer. It may lack the manufacturing capacity or the advanced materials needed for military or high-technology products. For military technology, there may be a lag of 5 to 10 years before the stolen technology appears in a foreign weapon system (e.g., stealth aircraft or submarine technology). The pace of adoption may be faster for some commercial technologies (high-speed trains, search engines, or wind power generators), while in other instances (semiconductors), the production of a competing product may be delayed indefinitely.

The same is not true for business information, which has immediate utility. It can involve the theft of preparatory material for business negotiations, contract data, "insider" information valuable for investment, and, in several cases, oil exploration data that allow a competitor to bid before or underbid for rights. Commercial espionage is nothing new, but it has reached a level where it poses a risk to the security of the United States and its allies.

The "Farewell Dossier" is a useful example of strategic collection programs and how to respond to them.[31] In the 1970s and 1980s, U.S. officials suspected an extensive, organized effort by the Soviet Union to illicitly acquire technology. However, there was only anecdotal evidence and it was difficult to focus policymakers on counterintelligence until an allied service recruited a Soviet officer, Vladimir Vetrov, who provided substantial evidence of a program whose scope, scale, and success had been unrecognized. We may now have a "Farewell" in cyberspace. In the absence of a Vetrov, can we "reverse-engineer" the collection programs of foreign opponents as a way to

[31] Guss W. Weiss, "The Farewell Dossier: Duping the Soviets," CSI Studies, April 14, 2007, https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/96unclass/farewell.htm.

highlight the importance of a comprehensive counterintelligence effort in response? As a starting point, we can ask how well do national collection programs correlate with national strategies for economic and technological development, and where these strategies may be useful for guiding counterintelligence efforts. Such estimates can help guide U.S. responses.

The Farewell incident is a useful (if not precise) precedent for creating consequences. Vetrov's information provided the names of over 200 Line X (the KGB division responsible) officers assigned to illicit technology in 10 KGB residencies, along with information about more than 100 Western informants. The Soviets were successful in meeting the majority of their collection goals for radar, semiconductors, and other industrial equipment.

Because Line X used traditional espionage techniques (officers under official cover who recruited agents at high-technology companies), the United States and NATO allies were able to disrupt its operations by suddenly expelling 200 officials. The United States also tightened controls on technology transfer and a deception program that provided the Soviets with defective components and plans.[32]

China's cyber espionage combines official programs and the coordination of unruly efforts of individuals, companies, and civil agencies as collectors. A broad range of proxies engage in cyber espionage at the behest of the state. This diffuse, cyber espionage collection program reflects China's approach to intelligence collection—instead of relying on officers operating under official cover, China's uses what been described as "a thousand grains of sand," where businessmen, researchers, and students are asked to collect information when they visit another country.[33]

Responding to Espionage

Vigorous response is the key to managing cyber espionage. It is normal, when one nation catches another's spy, to expel an attaché, recall an ambassador, call the spying nation's ambassador in for a harsh meeting, or develop other punitive measures,. Of course, the guilty parties will deny everything and demand to see the evidence. The normal response is to simply ignore these requests as de riguer public statements pronouncing innocence. The question of evidence sufficient to make U.S. officials comfortable with taking action has been a false obstacle in cybersecurity and, in any case, there are now so many incidents involving China that not directly addressing the problems seems peculiar. This is not a court of law, and lower evidentiary standards apply or the outcome will be paralysis.

One of the commonplaces of cybersecurity is that attribution is hard, if not impossible. The private sector relies on forensic techniques to examine code and identify the source of malware. Governments have many other sources. Attribution can be greatly increased if one uses "active measures" to gather intelligence on an opponent's cyber activities. This can include signal intelligence, human agents, and other techniques. It is the ability to fuse signals intelligence, cyber espionage, and traditional espionage by agents that gives intelligence agencies a superior ability to attribute cyber actions.

[32] Ibid. "Contrived computer chips found their way into Soviet military equipment, flawed turbines were installed on a gas pipeline, and defective plans disrupted the output of chemical plants and a tractor factory. The Pentagon introduced misleading information pertinent to stealth aircraft, space defense, and tactical aircraft."

[33] See, e.g., Northrop Grumman Corporation, "Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation," U.S.-China Economic and Security Review Commission, October 2009, http://www.uscc.gov/researchpapers/2009/NorthropGrumman_PRC_Cyber_Paper_-FINAL_Approved%20Report_16Oct2009.pdf.

National technical means[34] include satellites and other assets that are able to intercept or monitor a range of signals from the electromagnetic spectrum. Signals intelligence, which intercepts communications, is most important for cybersecurity. Cryptologic support allows advanced agencies to defeat encrypted malware. Massive data storage and computing resources allow agencies to correlate information from different sources, identify patterns, and establish attribution. If a target visits the attacking country and uses a mobile device, the intelligence agencies access to domestic telecommunications networks will allow it to implant malware. Most countries have some relationship and access to domestic telecommunications (for law enforcement or domestic security purposes) that allow them to do this.

The most direct form of intelligence support is to penetrate and access opponent networks, to hack the hacker. This can allow the identification of opponents' plans, capabilities, and intentions. You can see what the opponent is devising and read his strategies for using it. You can read his lists of proxies. Other covert identification techniques are available when the opponent network is accessed. The ability to monitor Internet traffic on a global basis allows the identification of patterns and sources. Success is never perfect, but the ability to correlate this information with other sources increases the chance of accurate identification.

Another important source involves human agents. Recruiting an agent who can provide details of an opponent's efforts and access to an opponent's networks reinforces any cyber of technical collection. Take the example of a secret program that develops malware that, from forensic methods, is absolutely unattributable. But if a government can get an agent into the program, they will have good knowledge of the source. Human agents can sprinkle infected thumb drives in parking lots or bathrooms, knowing that someone is likely to pick one up and insert it into a network (this technique was used against both the U.S. Central Command's classified networks and the Iranian nuclear program). An agent can access a target's laptop or mobile device and insert malware surreptitiously. Unattended devices in hotel rooms can be accessed and infected. Active intelligence reduces the attribution problem and may even allow, in the future, for a degree of verification of any agreement on constrains.

Identifying the author of a single event is difficult, but with multiple events, the likelihood of attribution increases. Attribution can be reinforced by active intelligence measures, using methods not available to private actors. The need for attribution varies by scenario. Attribution is a problem for law enforcement, with its high evidentiary standards. It is less of a problem in military conflict. There is always uncertainty in warfare, requiring judgments by commanders and policymakers; cyber warfare will be no different. Similarly, uncertainty is normal in espionage (which by its nature is covert), and decisions must be made using slight, incomplete, or unreliable information.

Other actions are also used to indicate displeasure, such as canceling official visits, freezing visas issuance, ending scientific cooperation, and, in extreme cases, imposing sanctions. An invitation by a senior official to the Dalai Lama (accompanied by a discreet warning to the Chinese that this is in response to unbridled espionage) or an expression of support, intangible or otherwise, can be used to respond.

The notion of responding by leaking, or threatening to leak, evidence that America has acquired though its own intelligence means on the corrupt practices of senior Chinese officials is unlikely to

[34] A phrase originally found in the Strategic Arms Limitation Talks that refers to technology-based intelligence activities.

be effective. Such leaks threaten the stability of the Chinese regime, a very disproportional response that could increase risk for the United States and destabilize the larger bilateral relationship. A discreet warning that unless the United States saw some reduction in cyber espionage is risky, but might work if the incident was grave enough to justify this action and if the individuals being threatened actually had control over the hackers. A threat to disclose corruption might easily backfire by provoking a damaging reaction, as it would not seem proportional.

Two incidents in addition to Farewell demonstrate the use by the United States of these and similar techniques in response to rampant foreign intelligence activities. In the mid-1990s, the then–director of central intelligence went to a European ally and suggested that it reduce its collection activities or face reprisals. It did so. In the early 1990s, the then–secretary of state gave a similar message to a Middle Eastern country. It did not heed his advice, and the United States took a number of steps to penalize it. That country also reduced its activities once it realized the seriousness with which the United States took its actions and brought to its warnings.

In each case, bilateral relations were damaged for many years afterward. U.S. expectations are shaped not only by its own practice, which is to not engage in illicitly acquiring business information or technology, but also by the practices of other nations and their willingness to heed warnings. The 2012 NCIX report is an indicator of the growing importance of this issue and of the potential for increased tension and potentially serious damage in the bilateral relationship. In deciding how to respond and which measure to use, the United States will need to consider the larger bilateral relationships (as it would in any intelligence matter).[35] Of course, this must be balanced against larger bilateral relations, but the United States' calculus for this consideration must use a more realistic appreciation of the damage to its economy and to its military strength from cyber espionage.

Cyber Espionage as an Economic Issue

One area of consequence that the United States has never pursued is in the WTO. The explanation for this is that the WTO processes themselves are very legalistic and that the United States for many years has lacked a strategic vision for international trade. Trade lawyers will tell you that in the absence of compelling evidence, there is no way to take cyber espionage and the theft of IP before the WTO.

This is far too timid. A state always has the right to exercise "force majeure," to go to a body and say that it agreed to procedures and concession on the grounds that other parties would similarly honor their commitments, and since they are not, the agreement no longer holds. This is a major step, but even to discuss it would raise the stakes for China and others to continue their camping of espionage and the theft of IP. The United States could probably also find major economic partners that would be willing to join it in this effort. Some would say that this would risk the collapse of the WTO, but a carefully managed discussion of repercussions for economic espionage could mitigate this risk. China and others might threaten to withdraw, but they are unlikely to do so as they benefit the most from the agreement. An astute diplomatic strategy would first pose the issues before taking any formal action. Even a credible hint that the United States is considering this would have an immediate effect on Chinese decisionmaking.

[35] Office of the National Counterintelligence Executive, "Report to the Congress on Foreign Intelligence collection and Industrial Espionage, 2009–2011," http://www.ncix.gov/publications/reports/fecie_all/Foreign_Economic_Collection_2011.pdf.

The WTO itself is not the right forum for dealing with national security issues, but compliance has led to a situation where the United States, which adheres more closely to the rules than its opponents, is at a disadvantage in responding to cyber espionage. A strategy for cybersecurity needs to restore flexibility to U.S. policy and reframe the debate in ways that offer solutions rather than paralysis. This dilemma reflects a larger problem with existing global institutions, including the WTO, which were created under very different political circumstances and with different expectations as to how members would behave. Whether China never intended to comply with its WTO commitments or whether it is unable to comply is irrelevant to the question of what actions can limit the loss of IP through cyber espionage. The WTO itself was not designed to deal with widespread noncompliance. A solution must include actions that are be external to WTO rules and procedures—hence the necessity of raising the possibility of force majeure.

WTO rules include security exceptions that clarify that the agreed should not be construed as preventing "a Member from taking any action which it considers necessary for the protection of its essential security interests." The exception provides the vehicle for the United States and its allies to move outside the constraints of the WTO to address cyber espionage. It would be a significant reinterpretation of this exception to use it to justify a vigorous response (or to threaten a vigorous response) to cyber espionage. Many in the trade community would oppose this kind of dramatic action, but a public discussion of it by U.S. officials and initial steps may be necessary to change the behavior of other states.

Those who would oppose such steps would include companies with significant interests in China that, with reason, fear retaliation. The United States could manage this by letting the Chinese government know that it would take appropriate countermeasures against Chinese firms. Given the symbiotic nature of the two nations' economic relationship, each side cannot go too far; but in this relationship, the United States is probably less vulnerable to Chinese pressure (even if individual U.S. companies are more vulnerable) than China is to U.S. pressure on market access. This is a case where the public good may outweigh the good of individual companies.[36] The preferred outcome, of course, would be a cooperative understanding with China on the moderation of its overly exuberant collection programs, but to achieve this goal the United States will need to take firm negotiating positions. The trade community would argue that the risk to the global institution is too great and that strengthening the WTO should be America's first priority. This view is too parochial to serve the national interest. The United States has significantly underplayed its hand in the cyber arena.

Consideration of whether to use a multilateral discussion of force majeure in the global trade system to address cyber espionage, and to follow such a discussion with actual restrictions on trade if discussion alone does not change opponents' behavior, is a significant foreign policy debate. An accurate assessment will require a view of international strategy that does not separate issues into the narrow and bounded channels preferred in Washington, of trade and security. Trade is a national security issue because it affects the sources of American power. A debate on force majeure and associated measures puts to the test the assertion about "death by a thousand cuts," or the "hollowing out of the American economy by cyber espionage." If this is a serious threat, it deserves serious actions. If we cannot find the political will to overcome parochial interests and take serious action, we must either accept the damage or reconsider the importance of the threat.

[36] This calculation requires a separate discussion of reliance on exports and technology transfer to show the extent of China dependence.

While there are implicit understandings among nations as to what level and kind of espionage activity is acceptable (if discovered), there are no formal international agreements on espionage because it is an activity where sovereigns refuse to be constrained. For cyber espionage, formal agreement to constrain cyber espionage would involve asymmetric trades, where each party would have to give up more than it gained. The cyber espionage problem may best be approached indirectly, as an aspect of trade and IP protection, or law enforcement cooperation against cybercrime, or as constraints on certain modes of cyberattack. The United States can buttress these actions, however, by the creation of norms for state behavior and by improvements to the "governance of global networks."

4 ENGAGEMENT, NEGOTIATION, AND STABILITY

It is political failure, not technology, which gives advantage to the attacker in cyber actions. Cyberspace is an ungoverned terrain for malicious action, and both the United States and its opponents have not been slow to exploit this. Nations are no more likely to renounce cyberattack than they are firearms. The widespread commercial availability, the black markets in malware, and disenchantment with existing governments mean that even if states renounce cyberattack, individuals and groups will still have access to attack techniques. It is best to prepare for this future.

Since nations will not renounce cyberattack capabilities, we cannot expect to eliminate the use of cyberattack (if only because cyber espionage techniques can easily and quickly be adapted for attack). Given the covert nature of most cyber programs, we cannot assume that any agreement will eliminate the risk of cyberattack. At some later time, perhaps, cyber technologies may become secure, but this point is distant, and even if technology improves, an opponent can always exploit human error—something that will never go away—to penetrate and attack. Cyberattack is something with which we will have to live, and this means the U.S. should focus on how best to manage and constrain risk to create a stable environment.

This is not a dystopian future, although if we fail to take action to reduce risk, it will become one. We have not taken the needed steps to build shared understandings among nations on the new military capabilities; nor have we replaced the amateurish governance structure, with one that better serves global needs. Most people are unaware of risk until it affects them directly, and the danger is that despite all the noise, cyberattack, like terrorism in the 1990s, will seem of little concern and a remote threat. Perhaps the only comfort is that almost all other nations share this weakness.

Diplomatic engagement is the principal tool for reshaping this environment. Creating a better international framework for cyber activities will make the cyber environment more stable and reduce the sense of risk that many nations feel. We can draw upon the experience of the past in arms control and nonproliferation but must realize their limitations in the new international security environment. A new style of security negotiations will go beyond the traditional boundaries of politico-military engagement. Some cyber threats can only be addressed through indirect action, using agreements on trade or law enforcement cooperation to restrain cyber espionage, the use of proxies, or nonstate actors.

Formal agreements to limit espionage are unworkable. This reflects the low level of trust that prevails in the cyber "domain," where covert action is the norm, verification is difficult, and the incentives to cheat are high. An offer to reduce U.S. cyberattack capabilities in exchange for a reduction in opponents' cyber espionage will not work because the value to an opponent—the potential benefit of an agreement that reduces the chances of a U.S. cyberattack in exchange for giving up the very tangible benefits of cyber espionage and asymmetric attack will be unappealing. The practice of espionage is already embedded in the informal understandings among nations that govern espionage. The best outcome internationally would be to develop means to avoid destabilizing effects form espionage mistaken as the precursor to attack and, for the U.S., to find ways to limit opponent espionage activities.

Understandings and agreements to shape and limit the use of cyberattacks are more feasible. A reasonable negotiating goal for the United States would be to increase international stability by creating shared understandings on how cyberattacks will be used and what rules apply to them. This would buttress opponents understandings of U.S. redlines and thresholds and reduce the chances of miscalculation. America can make the cyber domain more stable by reducing both the chance of miscalculation and misperception. Norms are useful in this because they shape international opinion and affect (in varying degrees) political decision by national leaders. Since norms need not take the form of binding agreements, it is easier to obtain multilateral agreement to them. In an environment marked by a high level of distrust on all sides, agreement to reduce risk may be incremental, following a sequence of first confidence building measures to create the trust necessary for agreement, then norms for state behavior, and finally, perhaps, some binding agreement.

Strategic stability, a concept from the Cold War, is a situation where adversaries understand that there is no advantage to be gained from attacks against the opponent's industrial base, centers of population, or military forces. The "strategic stability" of the Cold War was linked to deterrence, which meant that stability could be "operationalized" through arms acquisitions programs that maintained deterrent capabilities and a balance between opponents. Opponent planning could be manipulated by investments in strategic weapons programs. This will not work for cybersecurity. Strategic stability is possible in cyberspace, but on a more limited basis and using different techniques to achieve it. Strategic stability implies that while there will be continued competition and low-level actions, these will not rise to the level of armed conflict between major powers that threatens vital national interests. This kind of stability may be the best outcome one can achieve now for the politico-military aspects of cybersecurity.

It would also be useful engage other nations to describe and reach agreement on what sorts of cyber actions are more likely to be misinterpreted. While cyber attack and cyber espionage may look the same in their initial stages, there is no intelligence value to hacking an industrial control system. A shared understandings that action against these industrial control targets would be interpreted as the prelude to an attack could begin to set bounds for the use of cyberattack. A shared understanding on how a penetration of industrial control systems would be interpreted is more a precise and more implementable than declaring all critical infrastructure targets off limits. It recognizes that attacks on these targets will be attractive to attackers in many different conflict scenarios, and instead of relying upon opponent goodwill it establishes a boundary clear to both attacker and defender. .

We can envision a situation (as in traditional arms control) where opponents could negotiate restrain by agreeing to forgo some capacity in exchange for some matching limitation by an opponent. However, we are so far from the level of precision and openness in any multilateral or bilateral discussion of cyberattack needed for this level of agreement. Given the nature of cyberattack, we may never be able to reach agreement to limit certain kinds of "weapons."

Cold War arms control built on a shared understanding among opponents of the risks they faced from nuclear war. Cyber opponents also face the risk that unchecked malicious action will damage a global infrastructure upon which they depend, but they have not reached the point where they believe that that risk of damage to this global infrastructure outweighs the benefits of malicious action.

The experience of arms control offers some precedents that, with care, can be used in considering international cybersecurity. In the Cold War, initial, limited agreements let opponents develop enough trust to begin more serious talks. These initial agreements began in 1963, after the Cuban Missile Crisis, with a "hotline" and by the end of the decade included the Non-Proliferation Treaty, and agreement to limit the spread of nuclear weapons. Similar agreements for cybersecurity may need to begin with basic confidence-building measures such as crisis communications, and could eventually involve shared understandings on norms and thresholds, transparency, and self-imposed limits on action.

Unlike nonproliferation discussions, where like-minded nations cooperated to reduce threats to international security, arms control negotiations were discussions between opponents who sought to reduce the risk to themselves while perhaps gaining an advantage over the other. Binding agreements to control arms and to reduce risk were buttressed by crisis management mechanisms and by confidence-building measures that sought to reduce mistrust and suspicion. Ultimately, arms control negotiation required the United States to decide where it would improve its security by agreeing to forgo some capability in exchange for some concession from its opponent or where creating a new military capability better served national security.

Such trades for cybersecurity are limited, as most would be asymmetrical—the United States would give up more than it could reasonably expect to receive in return from an opponent. The worst outcome would be an arrangement where the United States gave up an area of advantage in exchange for an unverifiable commitment for restraint. We can set a number of threshold questions for assessing when it is in the U.S. interest to give up some capability in exchange for a promise from a potential opponent to show restraint:

- Does the United States have unique capabilities or do other nations have similar or equivalent cyberattack capabilities?
- If the United States does have a unique capability, does a potential opponent possess some other cyberattack capability that the United States does not for which it would be willing to "trade"?
- How likely are opponents to honor any commitment, and does the United States have the means to verify this?

Regarding unique capabilities, the best answer is that U.S. offensive cyber capabilities are better, but not unique when measured by the capability to cause disruption, and that this lead is eroding as others improve or acquire cyberattack capabilities. Similarly, any asymmetric advantage potential opponents had because the U.S. was more reliant on computer networks is rapidly disappearing. Our most likely state opponents are now sufficiently dependent on digital networks for military and economic activities that they present inviting and vulnerable targets.

The U.S. was quick to take advantage of cyberspace for military and intelligence purposes, and potential opponents have an exaggerated sense of U.S. military capabilities. This may inflate their expectations as to what the United States should be willing to give up. This means that any U.S. offer may be seen as insufficient or duplicitous. Changing this would require the United States to provide credible but very sensitive information that could compromise U.S. operational capabilities—the Cold War experience suggests that in some instances this inadvertent transfer can occur.

America's most likely opponents do have some unique capabilities that it has chosen not to match, in their use of proxies and (for China) the extensive economic espionage programs. So one potential trade would be for the United States to renounce some offensive capability it has in exchange for Russia and China to end their use of proxies. In any such agreement, however, Russia and China would be giving up ongoing activities from which they derive real benefit in exchange for a U.S. promise not to engage in some objectionable future activity by the U.S.

For this to be worthwhile, Russia and China would have to assess the likelihood of the U.S. carrying out that activity as very high. Since they are likely to believe that the United States would only engage in offensive cyber operations in the context of some larger conflict, they would also assess the likelihood of that conflict. If they decide that the United States is unlikely to launch offensive cyber actions against them outside a conflict, the trade does not make sense for them. They would be giving up an activity of immediate benefit in exchange for an U.S. promise to limit some future harm, in circumstances (armed conflict with the U.S.) that are themselves improbable.

The United States can affect this opponent calculation of the value of agreements on restraint by creating consequences (by itself or with others) for malicious cyber actions. It could help establish norms of state behavior that would penalize the use of proxies and economic cyber espionage. It could, through the release of doctrine and public statements, imply that the risk of it using offensive cyber actions is greater than opponents estimate. While the calculus of negotiation, the implicit formulas and weightings used by states to decide when an agreement serves their interests, currently does not favor agreement on restraint, this can be changed by U.S. action.

One standard strategy in strategic arms control was to begin to build new weapons to force the opponent into a more favorable negotiating position. This is unlikely to work for the United States in any negotiation on cyberattack. Opponents already fear the United States, and in any case, its most advanced capabilities must remain secret to maintain their effectiveness. It is also easy for opponents to match the United States in any building program. If there is a military rationale for the United States to create new attack techniques, they should be pursued, but we should not develop new techniques with the intention of using them a bargaining chip to win concession from opponents.

The strength of any binding agreement on restraint will depend upon the ability to verify it. Verification, as in traditional arms control, will require a reliance on the national technical means available to the United States and its close allies. For example, the United States could conduct cyber espionage against another signatory to determine if it was living up to an agreement to forgo certain malicious cyber activities. This could be reinforce by signals intelligence (the interception of telecommunications) and by the recruitment of agents.

This combination of cyber, SIGINT, and HUMINT sounds adequate, but in practice it may be difficult to be confident that we have adequate knowledge of an opponent's intentions and capabilities for cyberattack. National technical means directed against Soviet strategic forces could rely on monitoring Soviet test data from weapons tests and satellite imagery. These sources are largely unimpeachable, but they not useful in the cyber domain. Good technical collection still provides only imperfect information on opponents' cyberattack developments.

The ubiquitous nature of information technology and global connectivity paradoxically may make it easier to hide programs. Russian proxies include very skilled hackers with advanced knowledge of computers and networks. A cyberattack development effort could hide in a sea of commercial activity, disguised as cybersecurity research or some other peaceful or defensive activity. The multiplicity of potential targets for monitoring could overwhelm finite collection resources. The United States could attempt verification by national technical means, but we would not be able to have as much confidence in the results as we did for strategic arms during the Cold War.

Verification of compliance with a norm involves similar verification problems, but the "loss form noncompliance is lower as norms are usually not based on mutual concessions. A norm is "operationalized" because the adhering government chooses to do so, not because the other side has agreed to give up some military advantage. Violating or ignoring a norm may bring international opprobrium, but it does result in in the "cheating" side gaining an immediate advantage and the conceding side put immediately in a disadvantageous position. Verification difficulties need not stand in the way of agreement on norms.

The Limits of Disarmament

Should the United States accept additional constraints on the development and use of cyberattack techniques, such as a no-first-use policy, a ban on cyber weapons, or the creation of sanctuaries—targets that are off limits to attack? These ideas are staples of the disarmament agenda, but none of them make any sense when applied to cyberattack.

Comparisons between cyberattack and nuclear strategy are too imprecise to be of much use—cyberattacks and nuclear weapons are not the same.[37] But parallels between nuclear and cyber exist in narrow circumstances. One similarity involves the growing discussion of "disarmament." No first use, for example, is an inheritance from the nuclear debates. The intent of no first use is to increase stability by reducing the risk that an opponent may misinterpret actions as a first strike and react in what they think is a preemptive fashion. No first use depends on the ability to trust an opponent's pledge, reinforced by the international opprobrium that violating such a pledge would bring. Unfortunately, while America's likely opponents do not enjoy international opprobrium, it is not enough to change their actions in the near term or deter them from a course of action they deem vital. When a nation routinely engages in violations of human rights or support for oppressive regimes, this does not signal a great concern for the opinion of the international community. Nor does an unwillingness to accept national responsibilities in accordance with law and practice to control criminals resident on a national territory (because they are in fact proxy forces) build confidence in observance of a no-first-use pledge.

No first use is also a diplomatic ploy used in the past to constrain perceived areas of U.S. military advantage. China and Russia both used this ploy during the Cold War. Countries calculate that a no first use pledge "costs" them less in terms of military advantage than it costs the United States, particularly if they do not intend to abide by that pledge.

No first use makes little sense when it is taken out of the nuclear context. Nuclear weapons are so devastating that they justify additional constraints on their use that go beyond those found in international law. The central existing constraint is that the use of for force is justified only in self-defense (to preserve territorial integrity or political independence) or when authorized by the UN Security Council in the interest of international peace. If cyberattack is really a "conventional weapon" not much different from artillery or aircraft, if its use does not produce horror and devastation, it is reasonable to reject the idea that a special constraint like no first use goes beyond existing international law is necessary or helpful.

Existing international law on armed conflict governs and restricts the use of military force against civilian targets. It does not provide for sanctuaries where force is banned. This is a pragmatic acknowledgment of the realities of combat. Civilian targets can be attacked if there is an overriding military necessity (in the judgment of commanders) and if due concern is taken to minimize harm to civilians (through proportionality and discrimination). Identification of a particular target set as a sanctuary faces the risk that an opponent will exploit this for military advantage.

[37] Joseph S. Nye Jr., "Nuclear Lessons for Cyber Security," *Strategic Studies Quarterly* 5, no. 4 (Winter 2011), http://www.au.af.mil/au/ssq/2011/winter/winter11.pdf.

Perhaps more important, some of the new classes of opponents that America faces do not seem greatly interested in observing any restriction on civilian targets. The logic of jihadis allows them to strike targets that would, under Western practice, be considered prohibited. They would justify these actions on the grounds that Western military actions cause indiscriminate civilian casualties among Muslim populations or that or that there is a religious sanction for attacking unbelievers. Other potential opponents, such as Iran and North Korea, have never shown any hesitation in attacking civilian targets once they have concluded that it serves their self-interest to do so. It is not even clear that some classes of opponents accept the validity of international law on the use of force or are aware of its existence. In these circumstances, a policy of identifying certain classes of targets as sanctuaries is an invitation to cheat.

A review of Soviet planning in the Cold War show that it targeted Western European power grids, telecommunications services, transportation hubs, fuel pipelines, and government centers for strikes Including the use of WMD) on the opening day of any conflict. Publicly available documents on current Russian military doctrine suggest this has not changed. Chinese military doctrine is similar in considering civilian targets as legitimate. Whether either country (or any other nation possessing such capabilities) would choose to attack infrastructure is a different question whose answer would depend on circumstances and the course of military action, no opponent is likely to renounce the possibility of such attacks.

Guerrilla groups routinely attack a similar set of critical infrastructure targets—pulling down electric high-tension electrical wires or and telephone lines, or attacking government centers. These kinds of attacks are normal for insurgents and should a group ever acquire the ability to use cyberattacks, they will likely consider striking infrastructure. Similarly, strategic bombing campaigns targeted critical infrastructures as part of an air campaign to reduce an opponent's capacity and will to resist. Civilian critical infrastructure is a legitimate target, and attacking it is routine in warfare. We should not expect to successfully exclude cyberattack from this routine military activity.

Nor is the idea of preventing the "proliferation" of cyber weapons a useful strategy to reduce risk. Widespread commercial availability undercuts any effort at control. Cyberattacks use widely available commercial products—roughly 400 million desktops and notebook computers were produced in 2001, for example, and estimates place the number of such machines produced since 2004 at over 2.3 billion.[38] Software coding skills are ubiquitous and easily acquired, and many of the tools available for legitimate network analysis and administration can be used for attack purposes. While it might be possible to close down the flourishing back market in cyber exploits and criminal software, these are not the source of the most advanced techniques.

Cyber exploitation techniques are also widely and routinely used in criminal and espionage activities, and the line between a penetration for the purposes of exfiltrating information and a penetration for disruption network services and data is very thin. The notion of calling for an end to espionage is laughable, and since espionage will continue, nations will have access to the tools needed for certain kinds of cyberattack. The widespread availability of commercial products and software makes the idea of banning "cyber weapons" suspect and impossible to define in a meaningful way. Specific code like that used in Stuxnet could be called a weapon, but the components

[38] Aditya Shah and Abhinav Dalal, "The Global Laptop Industry," Systems Realization Laboratory, Georgia Institute of Technology, April 13, 2009, http://srl.gatech.edu/Members/ashah/laptop_industry_analysis_aditya_abhinav.pdf.

and “precursors” for the “weapon” are so widely available in legitimate markets around the globe that any ban could be easily circumvented.

It has been possible for the international community to agree to ban certain horrific weapons of mass destruction, such as biological or chemical weapons, but these bans depend on the moral force created by the repugnance such weapons induce in most people. Even then, there are instances of cheating, including charges by the United States that Russia systematically violated its commitments under the Biological Weapons Convention for 30 years. Such charges do not create confidence that a cyber weapons ban would not be even more likely to be evaded. Cyberattacks do not generate the same repugnance, and a ban on their development would be impossible to verify without some kind of intrusive global monitoring that would raise serious privacy and sovereignty concerns; and even then, a determined opponent could likely evade detection.

The appeal of a disarmament agenda lies in the understandable desire to eradicate war and the use of force. A series of major obstacles make it very unlikely, however, that disarmament is a realistic option for cyber security. “Cyber weapons” cannot be defined, much less banned. Classes of civilian targets are already afforded a degree of protection under existing international law, but experience shows that they cannot be reasonably declared to be outside consideration for attack. No first use of cyberattacks could be considered, but nations are unlikely to stop engaging in cyber espionage, which could be misinterpreted as an attack, and there may be some pressure to use cyberattacks in the opening phase of any conflict to disable opponents’ capabilities. The inadequacy and unimplementable nature of a disarmament approach means that cyber security must be sought elsewhere and through other measures.

We could treat cyberattack as if it was a weapon of mass destruction, whose effects are so horrific that there are few circumstances in which their use can be contemplated. The use of weapons of mass destruction is stigmatized. There are no circumstances where the governments, media, and influential individuals who make up the international community believe it is legitimate to use chemical or biological weapons. This stigmatization has led to formal bans on chemical and biological weapons. Nuclear weapons use is also subject to de facto stigmatization. There are only limited circumstances in which some nations say they will use nuclear weapons to defend their territory or allies. In the Cold War, the United States and the Soviets came to understand implicitly that nuclear weapons would be used only in extremis. The stigmatization of these weapons has led some to ask if a similar approach should be applied to cyber weapons. There are a number of reasons why this does not make sense.

The primary reason weapons are stigmatized is that they are horrific (in a sense, all weapons are horrific, but the scale and nature of the destruction caused by weapons of mass destruction puts them in a special category). Although the United States contemplated the use of nuclear weapons in the 1950s, with some in the military predicting that tactical weapons would soon be seen as “conventional,” the horror of nuclear war made the use of nuclear weapons politically unacceptable. The same is true for chemical and biological weapons. However, it is not true now or for the foreseeable future that cyber weapons will have a horrific effect. Efforts to develop various scenarios that equate the effect of a cyberattack to a weapon of mass destruction rely on exaggeration and implausible scenarios and coincidences, and are not credible. Cyberattack will be an unavoidable part of future military conflict. Any policy must take into account the reality that cyber weapons will be used.

The close linkage to espionage makes countries reluctant to discuss or even admit that they possess cyber capabilities, and this linkage also makes a “ban” on first use of dubious value. A no-first-use commitment could require countries to renounce cyber espionage—something they are unlikely to do. Since the techniques of cyberattack and cyber espionage are similar, asking for a commitment not to develop or use cyber tools for the penetration of opponents' networks is really asking for a commitment not to spy. A no-first-use commitment could even be destabilizing if a victim were to misinterpret a cyber espionage exploit as an attack.

If we treat cyberattacks like other conventional weapons, a set of rules involving proportionality, discrimination, and distinction apply to attacks and to targets. These principles imply that an attacker would need to assess the potential for collateral damage to civilian targets for a cyberattack to be lawful. The use of cyberattacks during conflict would face the same constraints as attacks using kinetic weapons. This approach may not satisfy proponents of disarmament, but it is more likely to allow risks to be controlled than a blanket prohibition that will inevitably be violated. The guiding principle is that he U.S. should accept no agreement that constrains it more than its opponents.

5 PRINCIPLES AND CONCLUSIONS

We cannot avoid the conclusion that the use of cyberattack is unavoidable in military conflict and that advanced militaries will prepare the means for cyberattack. Use of cyberattack in conflict will follow the same decision-making processes as other weapons. For "strategic" use of cyberattacks, determining factors will be the long range, speed, the difficulty of defense and the possibility of increased surprise that cyberattack provides and that was previously only afforded by ballistic missiles.

The use of cyberattacks outside conflict, and the use of cyber espionage, will be shaped by an attacker's belief in the likelihood of attribution and potential consequences if detected. The risk of cyberattack by nonstate actors is growing but may never match the capabilities of nation-state capabilities, given the disparity in resources. This nonstate risk is not amenable to the same tools that will mitigate the risk of military use. Cyber espionage and state-sponsored crime have become a normal part of online activity, in large part because of failures to establish bounds and cooperative measure on an international level, but the greatest risk from cybercrime and cyber espionage comes from miscalculation leading to escalation into more damaging military conflict.

Deterrence based on the threat of reprisal or the use of military force cannot be extended and does not work. It does not change opponents' calculations on use either in conflict or outside conflict. The most important means to shape opponents' calculations is to associate tangible consequences for malicious cyber actions. This is the only way to limit state-sponsored espionage and crime, which cannot be the subject of credible military threats.

Drawing on this discussion, the United States should base its approach to international cybersecurity on six principles:

- Cyberspace is not a unique environment. States will behave in this environment as they would in any other.
- We cannot "disarm" in cyberspace, and there will be no "global zero" for a cyberattack.
- We have entered a period of sustained, low-level competition for influence where opponents' miscalculations and misperceptions are a source of risk to the United States.
- U.S. interests are best served by embedding cyberattack and cyber espionage in the existing framework of international law, and long-term U.S. interests are best served by winning international agreement to this.
- America's immediate goal in negotiation should be to increase the risks of launching a cyberattack or engaging in malicious cyber activity for both state and nonstate opponents.
- There is a limit to what negotiation can achieve in reducing risk; there will always be risk. The U.S. goal should be to decrease and bound this risk as part of its larger efforts to strengthen international security.

Therefore, the United States should not accept any agreement to constrain a cyberattack. Instead, the goal for U.S. policy is to "normalize" the role and place of cyberspace in international relations and security, to create a more stable environment by reducing the chance of miscalculation, and by embedding the use of cyberattack and cyber espionage in existing framework of state relations.

Norms are an important part of this framework, but they are insufficient if not backed by action. The most important norms would establish state responsibility for actions in cyberspace that originate in their territory (whether state-sponsored or not—this is an extension of existing international practice) and the application of existing international law to cyberspace, and in particular the application of the laws of armed conflict. The core of an approach to norms is that cyberspace is not sui generis; nor is it a unique environment for international security.

Nations do not rush into agreements, even nonbinding agreements, that might impinge upon their sovereign rights. The calculus of how cybersecurity would affect these rights is not well established (as is the case with other security and economic issues). Claims that cyberattack is like nuclear warfare are spurious, but they create uncertainty that makes nations cautious in moving toward an agreement. Competing political agendas—one authoritarian and the other democratic—increase this caution. Agreement on norms will be difficult to achieve in the current international political environment; the widely recognized intermediate step is to get agreement on confidence-building measures to lay the groundwork for future agreement on norms.[39] The most important confidence-building measures involve transparency and the creation of internationally recognized "redlines." These are the most valuable for reducing misperception and miscalculation.

Any exchange of information at first will be asymmetric, with the United States and other democratic nations providing more than authoritarian regimes. The United States should emphasize reciprocity but should not expect it. A number of interim measures can build the case for reciprocity (or create the effect of mutual transparency, even if some nations do not participate). These could include official statements on how the United States perceives other nations' cyber doctrine and policies and what activities it considers to be problematic (as is done now with annual reports on human rights).

The core trade for international cybersecurity will involve a new model of governance in exchange for responsible behavior by states. The United States should link any change in governance to universal agreement on the acceptance of state responsibility. The old, multistakeholder model is inadequate and must be replaced (although not the by the overly governmental approaches suggested by authoritarian states). The United States and its partners can no longer treat Internet governance and cybersecurity as separate issues, that cyberspace is a unique domain not subject to the usual requirements of sovereignty. Technology has created a new domain for trade, discourse, and conflict. Millennial beliefs that existing "rules" could not apply to this domain have been tested and found wanting. A new approach to international security will recognize what is inevitable, extend international practice and law into cyberspace, and create a framework of goals, principles, and consequences that will create stability at a strategic level and best serve America.

[39] James A. Lewis and Katrina Timlin, *Cybersecurity and Cyberwarfare: Preliminary Assessment of National Doctrine and Organization* (Geneva: UNIDIR, 2011), http://www.unidir.org/pdf/ouvrages/pdf-1-92-9045-011-J-en.pdf.

ABOUT THE AUTHOR

JAMES ANDREW LEWIS is a senior fellow and director of the Technology and Public Policy Program at CSIS. Before joining CSIS, he worked at the Departments of State and Commerce as a Foreign Service officer and as a member of the Senior Executive Service. His current research examines the political effect of the Internet, strategic competition among nations, and technological innovation. Lewis received his Ph.D. from the University of Chicago.

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

1800 K Street, NW | Washington, DC 20006
Tel: (202) 887-0200 | Fax: (202) 775-3199
E-mail: books@csis.org | Web: www.csis.org