



SSQ

STRATEGIC STUDIES QUARTERLY

FALL 2012

VOL. 6, NO. 3

CYBER
SPECIAL EDITION

Commentaries

America's Air Force: Strong, Indispensable, and
Ready for the Twenty-First Century

Gen Norton A. Schwartz, USAF, Retired

Lt Col Teera Tony Tunyavongs, USAF

Claiming the Lost Cyber Heritage

Jason Healey

Depleted Trust in the Cyber Commons

Roger Hurwitz

Escalation Dynamics and Conflict Termination in
Cyberspace

Herbert Lin

Sharing the Cyber Journey

Maj Gen Suzanne M. Vautrinot, USAF

The Specter of Non-Obvious Warfare

Martin C. Libicki

Internet Governance and National Security

Panayotis A. Yannakogeorgos

The Customary International Law of Cyberspace

Col Gary Brown, USAF

Maj Keira Poellet, USAF



STRATEGIC STUDIES QUARTERLY

*An Air Force–Sponsored Strategic Forum on
National and International Security*

VOLUME 6

FALL 2012

NUMBER 3

Commentaries

- America's Air Force: Strong, Indispensable, and
Ready for the Twenty-First Century* 3
Gen Norton A. Schwartz, USAF, Retired
Lt Col Teera Tony Tunyavongs, USAF
- Claiming the Lost Cyber Heritage* 11
Jason Healey

Part I

Feature Article

- Depleted Trust in the Cyber Commons* 20
Roger Hurwitz

Perspectives

- Escalation Dynamics and Conflict Termination
in Cyberspace* 46
Herbert Lin
- Sharing the Cyber Journey* 71
Maj Gen Suzanne M. Vautrinot, USAF
- The Specter of Non-Obvious Warfare* 88
Martin C. Libicki
- Internet Governance and National Security* 102
Panayotis A. Yannakogeorgos
- The Customary International Law of Cyberspace* 126
Col Gary Brown, USAF
Maj Keira Poellet, USAF

Book Reviews

<i>Critical Code: Software Producibility for Defense</i>	146
National Research Council	
Reviewed by: Lt Col Deborah Dusek, USAF	
<i>Airpower for Strategic Effect</i>	147
Colin S. Gray	
Reviewed by: Benjamin S. Lambeth, PhD	
<i>Chinese Aerospace Power: Evolving Maritime Roles</i>	149
Edited by: Andrew S. Erickson and Lyle J. Goldstein	
Reviewed by: Capt Paul A. Stempel, USAF	

Part II (online only)

Cyber Power, National Security, and Collective Action in Cyberspace, 10-11 October 2012

AFRI Cyber Power Conference Proceedings—Online

Topics Include:

How can strategists more effectively confront the challenges of the cyber environment to understand the key principles of the domain?

What is the relationship between cyberspace, its usage, and adaptation for national security purposes and the socioeconomic forces shaping its character that could impact the Air Force/national security community mission over the next five years?

What are the best cyberspace approaches for using to influence perceptions of international actors for global and regional stability?

How can we reduce the stigmatization of cyber weapons and cyber attack?

Available early 2013 at <http://www.au.af.mil/au/ssq/>

For conference registration, see page 152.

America's Air Force

Strong, Indispensable, and Ready for the Twenty-First Century

After examining every aspect of the American effort in World War II, President Harry S. Truman and his military leadership team were convinced that the nation needed an independent military service to operate exclusively in the air domain. The legendary exploits of the US Army Air Forces in World War II demonstrated that airpower, through gaining and sustaining air superiority and providing close air support to ground forces, was a *sine qua non* for success in major land operations. Moreover, the Army Air Forces' achievements established that air forces, through providing airlift, reconnaissance-based intelligence, and strategic bombing, could create important effects that were largely independent of tactical support and, in fact, could affect all levels of conflict, oftentimes simultaneously. These Army Air Forces contributions that were so valuable to the Allied victory are the very ones that today, seven decades after the end of World War II, still provide a shared identity and sense of purpose for Airmen, and make the US Air Force critical to the national defense.

Raison d'être—Then, Now, and Tomorrow

As it was then, the ability of airpower today to produce significant operational outcomes requires its comprehensive and integrated employment. The US Air Force is able to employ airpower in this fashion—to strategic effect—because Airmen comprehend and appreciate airpower's rapidity, global range, versatility to conduct a variety of missions, and flexibility to produce outcomes at multiple levels.¹ Over the past 65 years, Airmen have refined their understanding of these attributes and therefore of their role as the nation's principal airpower provider.

Today, only the US Air Force leverages globally scaled yet regionally tailorable air, space, and cyber capabilities specifically to affect outcomes that are distinct from only the effective tactical support of surface forces. To be sure, Army aviation continues to support ground maneuvers, Navy aviation remains critical to the security of our maritime fleets on the open seas and in littoral operations, and Marine aviation continues to be

integral to expeditionary amphibious and Marine air-ground task force operations in support of littoral campaigns. And most certainly, Air Force airpower remains ever dependable in providing tactical support whenever and wherever it is needed.

But strategically oriented airpower—that which provides *Global Vigilance*, *Global Reach*, and *Global Power* with unrivaled speed, versatility, and flexibility—is nearly exclusive to the US Air Force, and will remain in decidedly high demand, as the latest defense strategic guidance predicts in enumerating the 10 primary mission areas of the US armed forces.² Many of these areas emphasize Air Force capabilities—for example: deterring and defeating aggression, projecting power in anti-access and area denial environments, conducting space and cyber operations, and maintaining the preponderance of our nation’s nuclear deterrent.

To fulfill these airpower-intensive mission areas, and to ensure requisite access to increasingly contested air and space domains, the nation will continue to need an air force—the US Air Force—that, in addition to ensuring continued timely, precise, and reliable support to its surface force teammates, is singularly dedicated to fulfilling the nation’s full-spectrum airpower needs. Steeped in a mindset that views the battlespace in all three dimensions, Airmen are conceptually unbounded by topographical features. The Air Force will continue to leverage the inherent characteristics of the entire expanse above the earth’s surface in order to provide the full spectrum of airborne capabilities, from close air support to air mobility to global strike.

It is with this perspective that Airmen instinctively unfurl the entire map of the battlespace, to gain greater situational awareness over a broader expanse of distance and time. To every Airman, emphasizing approaches that traverse “over” or “around” rather than “through” is the prevailing *modus operandi*. The Air Force is a service that operates with a holistic view of air and space, providing harmonized, seamless capabilities across the full spectrum of operations, even as surface activities necessarily transition between *terra firma* and the maritime.

However, to the casual observer, it would appear that the Air Force has been less involved, or possibly less relevant, in the nation’s post-9/11 pursuit of its adversaries. Perhaps this is understandable, given the ground-centric nature of the conflict and the sterling professionalism and performance of our supremely skilled Army, Marine Corps, and special

operations teammates during sustained operations in Iraq and Afghanistan. However, as we demonstrate below, this is not the complete story.

Still others have come to believe, mistakenly, that the adaptations that the Air Force prudently made during the past decade—adjustments that were necessary given the wartime challenges that we faced—have distracted Airmen from their enduring and core contributions.³ Quite the contrary, we Airmen in fact have focused on our enduring airpower contributions, even as we tended to a few noteworthy but nontraditional assignments, such as convoy and base security, and Provincial Reconstruction Team command opportunities. Other than addressing these and a few other exigencies, we Airmen have concentrated on what a first-rate independent air force is expected to provide for the nation that it serves. In the case of the US Air Force, it is those enduring contributions—control of the air and space domains; global intelligence, surveillance, and reconnaissance (ISR); rapid global mobility; and global strike—that Airmen have provided proudly and reliably since the establishment of the nation's independent air service.

In so doing, the Air Force not only has demonstrated its efficacy. It also has made the case that support roles and independent roles are not mutually exclusive, but rather reciprocally supportive. This is true particularly in modern warfare, which is becoming ever more interdependent across the various domains. For example, prior to Operation Desert Storm, artillery was arguably the most destructive force on the battlefield. Thereafter, surface forces have depended largely on airpower to destroy opposing forces, while air forces often count on ground forces to compel adversaries to abandon hardened or otherwise safer positions and to hazard into areas where they subsequently are more vulnerable to attack from above.

In this vein of increased interoperation, Air Force contributions in the last decade have been critical to enhanced and more meaningful integration across the military services and their primary operational domains—a point that is even more noteworthy considering that budgets of late have encouraged parochial retrenchment and protection of narrower institutional imperatives. Notable examples of contributions that have enhanced our integration and interoperation include

- advancing the state of air mobility with capabilities such as the Joint Precision Air Drop System;⁴

- expanding our ISR enterprise capacity to process, exploit, and disseminate timely, accurate, and relevant intelligence to tactical forces;
- assisting in revisions to close air support and joint fires doctrine to strengthen protection of friendly ground forces; and
- modifying aircraft and weapon systems such as the B-52 Stratofortress and B-1 Lancer to employ in new and innovative ways.

Representing our team-oriented approach, these innovations do not diminish our commitment to our core service contributions. Quite the opposite, these adaptations and others in fact have helped us to reconnect with our heritage while consistently helping to ensure our national defense. This reclaimed heritage has solidified confidence in our enduring functions and our ability to perform our duties well. We celebrate the many important ways in which Airmen have contributed and will continue to contribute to our nation's security and to fulfilling our geostrategic interests.

We find that our ground-force teammates have provided some of the most full-throated, wholehearted, and significant arguments for the efficacy, value, and reliability of the US Air Force. Throughout operations in Iraq, Afghanistan, and around the world, Airmen have refined their role according to operational requirements and urgent needs of the combatant commanders. In so doing, Airmen have become even more credible, dependable, and valued members of the joint team. Few, if any, division commanders would want to “go downtown” without Air Force bombers and fighters preparing and securing the battlespace.⁵ Hardly any company commander would unnecessarily hazard an enemy engagement without Air Force close air support. And, almost certainly, no platoon leader would prefer to guess what danger might be lurking around the corner, over the wall, or on the roof, rather than be with situational awareness par excellence from Air Force remotely piloted aircraft and their ability to target, track, and in many cases, engage the enemy.

Air Force Contributions to the Nation's Strategic Interests

Put another way, the Air Force is held in high regard by those who depend on its distinct capabilities the most. The fact that land warfare, by necessity, has been the US military's emphasis in the last decade does

not obviate the continued demand for strategically oriented and globally postured Air Force airpower. Indeed, this need will come into greater focus as the nation rebalances its strategic emphasis and effort toward the Asia- and Indo-Pacific. Accompanying this recalibration is the immediate challenge of substantially increased distance and time, both from the homeland and within the region itself, which covers 13 time zones and more than 100 million square miles.

It therefore is entirely clear that the nation will continue to depend on inherent airpower characteristics and unique Air Force contributions. Domain control, ISR, rapid global mobility, and global strike, as well as the additional distinctive ability to conduct high-volume, cross-domain command and control of air, space, and cyber capabilities, will remain essential to the nation's strategic interests. Essentially, this "four-plus-one" construct represents, most fundamentally, those capabilities and contributions that are at the core of the world's preeminent air force.

But Air Force contributions are valuable not only to the portfolio of US armed forces capabilities alone. The assured access to international airspace that the US Air Force provides is of tremendous importance to civil and commercial aviation as well. The United States, by many measures, is still the world's only genuine air and space nation, with strategic interests across its many dimensions—commercial, financial, diplomatic, legal, military, and others—that remain undeniably connected to aviation and aerospace. For example, the nation's economic health and prosperity are tied to the more than two billion passengers and some 35 percent of international trade (by value) that transit via international airspace annually. And, according to Federal Aviation Administration forecasts, air system capacity in "available seat miles"—the overall measure of commercial airline activity level, both domestically and internationally—will increase around 4.5 percent this year, and is anticipated to grow through 2031 at an average annual rate of 3.6 percent.⁶ These are but a few high-level statistics that presage a continuing upward trend in aviation and airpower's importance to our Nation's strategic interests. The US Air Force is prepared to maintain its place among the elite of the aerospace community, which has underpinned America's global awareness and influence since the early 20th century, and which will continue to leverage the advantages of air and space power for national effect in the 21st. However, with the proliferation of advanced technologies and high-speed computing that enable nonstate actors to exert influence in what formerly was the

exclusive domain of well-resourced nation-states, we must contend with a broader array of threats, including to the global commons. Among these threats are burgeoning anti-access and area denial challenges to our nation's ability to project global power, and competition in vital air and sea lanes of communication and transit that could turn unimpeded thoroughfares into crippled chokepoints. The US Air Force stands ready to meet these wide-ranging security challenges.⁷

The Air Force is prepared as well to continue providing our national leaders with strategic options that otherwise might not be available. Exemplifying this strategic versatility, flexibility, and readiness are the simultaneous operations of March 2011, when the Air Force, along with joint and coalition partners, spanned both intercontinental distances and the full continuum of operations to provide humanitarian relief in Japan and combat airpower and air support in Libya, all the while sustaining operations in Afghanistan and Iraq.

Conclusion

The four distinct Air Force contributions of control of the air and space domains; global intelligence, surveillance, and reconnaissance; rapid global mobility; and global strike represent not only our traditional core mission areas, but also those unique capabilities that will endure for the foreseeable future. They also serve as an anchor point around which all Airmen can rally with a core identity and a shared sense of purpose. Leveraging the inherent characteristics of air, space, and cyberspace into our unique and enduring contributions will be vital to our national interests in the future security environment. From potential higher-end conflict with near-peer competitors, to insurgencies and other localized and geographically distributed crises, to natural disasters and humanitarian crises—the need for airpower and its distinct advantages will endure.

The US Air Force is a proud and reliable member of the joint team. To face a future that will present wide-ranging challenges, we will have to leverage each unique strength within each of the military services. Every carefully tailored and considered contribution, bringing the capabilities of each and every military branch, is indispensable to the success of the joint team. Without the US Air Force working with its joint team members, there would not be a US armed force as we currently know it—certainly not one that can maintain its place as the most respected military in the world.

It therefore is ever more important that Airmen reaffirm and recommit to the core Air Force identity that gave rise to the nation's independent air service. For a service that has a heritage so closely tied to the advancement of technology, a deep appreciation for the key and enduring Air Force contributions is particularly important. This awareness strengthens us and allows us to adapt accordingly, as technologies advance, operational requirements emerge, and methods of warfare evolve. What once was primarily the domain of aviators is now necessarily trending toward greater prominence for operations other than manned flight—to name a few: space, remotely-piloted, and cyber operations—as well as the vital functions that battle-field Airmen perform “outside the wire,” shoulder-to-shoulder with their ground-force teammates. As the Air Force evolves according to changing domestic circumstances and dynamic global complexities, Airmen will find such diversity to be critical to the vitality of the Air Force. But we will remain as Airmen who have a clear appreciation for the core and enduring contributions, and the *raison d'être*, of the US Air Force.

Gen Norton A. Schwartz, USAF, Retired
Nineteenth USAF Chief of Staff

Lt Col Teera Tony Tunyavongs, USAF
*USAF Chief of Staff PhD Fellow, Fletcher School
of Law and Diplomacy, Tufts University*

Notes

1. For our purposes here, we are adopting the esoteric distinction between “versatility” and “flexibility.” See Air Force Doctrine Document 1, *Air Force Basic Doctrine, Organization, and Command*, October 14, 2011, 39–40.

2. *Defense Strategic Guidance*, 5 January 2012, 4–6.

3. See, for example, David W. Barno et al., *Sustainable Pre-eminence: Reforming the US Military at a Time of Strategic Change* (Washington: Center for a New American Security, May 2012), 43.

4. By raising the altitude for releases from 1,000 feet or less to 20,000 feet or more, JPADS kept aircrews safer and on more efficient flight profiles, while, in reducing the number of required convoys, it limited exposure of convoy personnel to what otherwise would be hazards in hostile zones. All told, JPADS improved accuracy and effectiveness in delivering food, cargo, and other vital supplies and materiel during the height of operations in Afghanistan and Iraq.

5. Norton A. Schwartz and Robert B. Stephan, "Don't Go Downtown without Us: The Role of Aerospace Power in Joint Urban Operations," *Aerospace Power Journal* 14, no. 1 (Spring 2000): 3–11.
6. *FAA Aerospace Forecast: Fiscal Years 2011–2031* (Washington: Department of Transportation, 2011), 5.

Claiming the Lost Cyber Heritage

The Air Force ensures that newer generations of Airmen learn through the vicarious experiences of those who have gone before them. They are taught to admire Eddie Rickenbacker and Billy Mitchell, and cadets and officers are tested to ensure they understand the lessons from Big Week, MiG Alley, and Rolling Thunder to Iraqi Freedom. Understanding this history and heritage is the primary way to turn the vicarious experiences of past generations into cumulative knowledge to educate Airmen of the future. According to the official Air Force website, heritage is “dedicated to the former Airmen who developed the independent Air Force and continue its evolution into cyberspace. . . . The people, events and equipment of the past are integral to understanding the future.”¹ Yet there is a particular heritage that has been forgotten and ignored as irrelevant. A recent search for “cyber” on official historical sites of the Air Force led to only four documents, no images, and a single video from 2012.²

Indeed, a fighter pilot that had never heard of the “hat in the ring”—who in fact spurned the history of airpower—would be an outcast. Yet this is not far from how the Air Force, and indeed the entire Department of Defense, treats the history of cyber conflict. Few, if any, Airmen involved in cyber operations today are likely to remember the major cyber conflicts, pioneering cyber leaders, doctrine, or units of the past.

How many of today’s Air Force cyber warriors know they can trace their lineage to AF cyber operations in the mid 1980s? Nearly 25 years ago a lone special agent in the Office of Special Investigations was intrigued by a call from an astronomer turned system administrator who found intruders in his networks at a national laboratory. The Air Force helped unravel an international espionage ring, nicknamed the Cuckoo’s Egg, where German hackers sought classified material on the Strategic Defense Initiative, which they sold to the Soviet KGB. Special Agent Jim Christy, the first cyber “ace,” is now retired but still delivering for the Air Force at the Defense Cyber Crime Center.

How many of today’s Air Force cyber warriors know when the Air Force declared cyberspace a new domain for military operations? The answer is not 2011 when the Department of Defense declared that the military would “treat cyberspace as an operational domain,” nor even in 2005 when the Air Force added cyberspace to its mission statement as a

domain in which to “fly, fight, and win,” but a decade earlier. In 1995 the secretary and chief of staff jointly signed the Foundations of Information Warfare which laid out basic definitions and principals for how the Air Force would work in cyberspace.

Before the Wright Brothers, air (while it obviously existed) was not a realm suitable for practical, widespread military operations. Similarly, information existed before the information age, but the information age changed the information realm’s characteristics so that widespread operations became practical.³ This statement is at least as good as anything written since by any military anywhere.

How many of today’s Air Force cyber warriors have even heard of the world’s first combat cyber unit? In 1996, the Air Force established the 609th Information Warfare Squadron (motto: “Anticipate or Perish”) at Shaw AFB to support CENTAF with combined offensive and defensive cyber missions “to fully operationalize information warfare on behalf of the JFACC [joint force air component commander] and the fighting forces.”⁴ This unit, the first such unit in the Air Force, is likely the first anywhere in the US military and the world.⁵ The unit invented the first INFOCON, now a standard defensive alert condition. It exercised heavily with CENTAF and “had control of the blue force air tasking order. They gave us a two-hour window to play in, and we got it within two hours,” according to the unit’s commander, then-lieutenant colonel Walter “Dusty” Rhoads, another Air Force cyber pioneer who had roles in every major joint cyber war-fighting organization for the next 10 years.⁶

These efforts at the 609th were just one part of using cyber to support the war fighter. As Maj Gen John Casciano, then head of AF intelligence put it in 1996,

Anything we do in the Air Force has to be consistent with a . . . JTF commander’s requirements and must meet those objectives. We believe that IW is absolutely critical and integral to Air Force operations at the JFACC level and below. We have some things to offer other communities, but our focus is on the operational and tactical levels of warfare. A lot of the targets and a lot of the things we would want to affect—command and control nodes and the adversary’s integrated air defense system (IADS)—are things the Air Force worries about on the battlefield.

How many of today’s Air Force cyber warriors know the first joint cyber commander was from the Air Force? It was not GEN Keith Alexander, USA, who took charge of US Cyber Command in 2010, but Maj Gen John “Soup” Campbell, USAF, the founding commander of the Joint Task

Force–Computer Network Defense in 1998. His approach to cyber operations was rooted deeply in his Air Force identity, “I grew up as a fighter pilot. My job was to blow things up, make smoking holes . . . so I always took it in that direction.”⁷

These are not empty facts or trivia for cyber operators to play on a long nightshift.⁸ They are emblematic of the rich heritage of the Air Force in cyberspace and illustrate the importance of learning the lessons of history. The Air Force is not responsible for all the problems of the Department of Defense in cyberspace. But it can fix those that it controls. If the Air Force is going to become the premiere force to fly, fight, and win in cyberspace, it must reclaim its proud cyber heritage and build “cyber-mindedness,” just as it has a tradition of air-mindedness. If it can succeed in this, the Air Force can again be seen as the cyber thought leaders in the military service and show the way for the other services, the Department of Defense, and the intelligence community. If not, the service is likely to continue to relearn old lessons and struggle under misperceptions with little relation to past experience.

Over two decades, the Air Force, and the Department of Defense in general, have made little progress on important policy and operational issues, but few realize just how little progress because few know how far back the story goes. For example, the sentiment behind the next two paragraphs should be familiar to many of today’s AF cyber professionals:

Nobody knew what a “cyber warrior” was by definition. It was a combination of past war fighters, J-3 types, a lot of communications people and a smattering of intelligence and planning people. . . .

The unfortunate part . . . was that the offensive side was still classified. You couldn’t even discuss it in an open forum. . . . But behind the scenes [we were] getting it integrated into the war fighters’ mentality, understanding the air tasking orders. . . . [We were] an Air Force unit and we had to understand how to get cyber introduced into the thinking of the commanders.

Unfortunately these quotes resemble those of today, but they are actually from Rhoads speaking about the 609th IWS in 1995. Likewise, consider the following quotes. One is from Rhoads, circa 1996, the other from Maj Gen Richard Webber of Twenty-fourth Air Force in 2009. *Why can’t we even tell the difference?*

I liken it to the very first aero squadron when they started with biplanes. We’re at the threshold of a new era. . . . We are not exactly sure how combat in this new dimension of cyberspace will unfold. We only know that we are the beginning.⁹

I almost feel like it's the early days of flight with the Wright Brothers. First of all you need to kind of figure out that domain, and how are we going to operate and maintain within that domain. So I think it will take a period of time and it's going to be growing.¹⁰

American Airmen learned how to dominate the aerial domain and deliver integrated combat effects in just 15 years between the first flight of the Wright Brothers and the Battle of Saint-Mihiel. Yet in the same amount of time since the first AF combat cyber unit, we have made so little progress in the cyber domain that quotes from key commanders a decade apart are indistinguishable.

This blindness to history has immediate operational implications. Much of what is treated as received wisdom is in fact not rooted at all in the history of cyber conflicts. Many of today's cyber warriors will tell you with all confidence that (1) cyber conflict is new and ever changing, (2) massive surprise attacks can easily prostrate nations, and (3) everything that is important happens at the speed of light. In fact, a study of cyber conflict history by the Atlantic Council and the Cyber Conflict Studies Association has shown that all three of these are incorrect or misleading.

There has been no essential discontinuity between cyber conflicts of 20 years ago and those of today. Of course, there are differences: adversaries have become more capable, underlying technologies (offensive and defensive) have changed, and corporations are now feeling the brunt of major espionage attacks. Yet, despite these developments, the dynamics of today's conflict would be familiar to the Airmen that fought them at the 609th Information Warfare Squadron in 1995.

Likewise, disruptive cyber attacks have so far tended to have effects that are either widespread but fleeting or persistent but narrowly focused. Few, if any, attacks so far have been both widespread and persistent. As with airpower, cyber attacks can easily take down many targets, but keeping many down over time has so far been out of the range of all but the most dangerous adversaries.¹¹

And strategically meaningful cyber conflicts rarely occur at the "speed of light" or at "network speed." True, individual tactical engagements can happen as quickly as our adversaries can click the Enter key, but cyber conflicts, such as Estonia, Georgia, Stuxnet, and the Conficker worm, are campaigns that take weeks, months, or even years of hostile contact between adversaries.

At least once before, the Air Force suffered similar “doctrinal lock in,” ignoring the emerging lessons from experiences in a new domain. In the 1930s, as all Airmen know, bomber enthusiasts preached that “the bomber would always get through,” across international borders and distances, and that hitting 154 known targets would quickly knock Germany out of the fight in six months.¹² Their exercises reflected this view, which left them completely unprepared for the lengthy attrition battles of World War II. The Army Air Corps lost nearly 10,000 bombers and took years to achieve strategic effects, having entered the war lacking appropriate doctrine, defensive firepower, and intelligence for targeting and bomb damage assessment.

Airmen learned that finding the right target for strategic effect is difficult, and there is a tremendous difference between temporarily disabling a target and permanently destroying it. Even with strategic attack in its DNA and a decades-long history of cyber conflict, the Air Force is still not recognizing the right lessons, much less learning them. It should be natural for the Air Force to realize that the “speed of light” of cyber operations is deceptive. There is no reason why Airmen should be fooled on this point, because they understand even though a dogfight can be over before the losing pilot even knows it has begun, an air campaign is rarely decided by a single tactical engagement.

By thinking only of conflict at the speed of light, the Air Force will overinvest in capabilities and doctrine to automatically counterattack and will be unprepared for the long cyber campaign most of our adversaries seem to expect and appreciate. If speed is mistakenly seen as the most important factor, then rules of engagement will allow ever lower levels to shoot back without seeking authorization—a relaxation of the rules, which may not be in the long-term economic or military interest of the United States. The Air Force will continue to dogfight blindly, flying from tactical engagement to tactical engagement without having thought about tomorrow’s battle or the one a year from now.

Similarly, Airmen should be the first to doubt it will be easy to have a prolonged strategic effect, even in cyberspace. If Flying Fortresses and Lancasters had difficulty achieving a strategic effect after dropping millions of tons of high explosives, we should never believe the fallacy that a few young hackers might take down the United States from their basement. This might be true in the movies or an espionage novel, but not in real life.

Yet basement-originated strategic warfare is a common theme from some who feel deterrence is difficult, since “cyberspace is fundamentally different. For someone with the right brainpower and the right cyber abilities, a cheap laptop and Internet connection is all it takes to be a major player in the domain.”¹³ These tools might help an adversary steal data or identities—even conduct a major intrusion like Solar Sunrise—but they are not sufficient to create a strategic effect requiring Air Force deterrent power.

This has been well known by Airmen since at least 1998 when Maj Gregory Rattray wrote his doctoral thesis, later published as *Strategic Warfare in Cyberspace*, with an extended comparison of how the early Army Air Corps struggles to learn how to fight in a new domain were directly comparable to what the Air Force was, and sadly still is, going through for cyberspace.¹⁴

These are all common misconceptions, but they are not supported by either the facts of cyber history or the experiences of Airmen. Perhaps soon, the world will see these kinds of attacks, but that is still no reason to ignore the past. By developing cyber-mindedness—a collective sense of the history, dynamics, possibilities, and limitations of cyber conflict—the Air Force can learn these and other critical lessons and prepare for the conflicts of the future.

The US Air Force has a longer, more distinguished heritage in the cyber domain than any other military in the world, but it is just one of the military services and should not be the *only* cyber service. As Major General Casciano put it in 1996 when he ran the AF cyber units, “We don’t claim [cyber] exclusively. We think we’ve got good ideas. We think we’ve got good capabilities. And we are reaching out to the other services and the joint community to offer what we have.”¹⁵ Fifteen years ago, this mindset helped the Air Force to be the world’s preeminent cyber force, but not anymore. “For a brief period,” as described by Lt Gen Bob Elder, retired, another AF cyber commander, “the AF was recognized as the thought leader on cyberspace, but when we narrowed our view, we undercut the basis for our leadership role.”¹⁶ Now retired, Major General Casciano echoes this sentiment, believing that “we have attempted to solve things organizationally and politically, not operationally.”¹⁷

To reclaim this heritage, there are a number of entirely practical steps the Air Force must take.

- Commission the Air Force Historical Research Agency to conduct oral histories of the pioneers of the Air Force cyber mission and collect

the official unit histories. This material should be the basis of a major study with appropriate lessons.

- Integrate cyber heritage and lessons into all professional military education (PME), starting with basic training and material for officer candidates (such as the Contrails guide) and continuing through all PME courses.
- Incorporate more detailed material on cyber heritage and lessons into classes such as Cyber 200 and 300 for the service's new cyber cadre.
- Encourage PME students to research and write on cyber heritage and lessons.
- Create a formal network to connect former AF cyber leaders, especially those retired or in the private sector. The Air Force created the earliest generation of cyber leaders, and many would enjoy the honor of being able to continue their association.

To further propagate this agenda, the Air Force Association—the main culture carrier for the service—is working with the Atlantic Council and the Cyber Conflict Studies Association to establish a distinguished panel of former AF leaders and cyber professionals to discuss other ways to build cyber mindedness and make the most of the service's cyber heritage. Some initiatives this group might consider may sound outlandish but are entirely reasonable if the Air Force indeed wants to establish itself as a force to “fly, fight, and win in air, space, and cyberspace.” These include:

- How might AF units earn battle streamers for participation in major cyber conflicts? For example, the AF Computer Emergency Response Team played significant roles in Solar Sunrise, Moonlight Maze, and Buckshot Yankee. These conflicts may or may not be sufficiently intense to qualify for a streamer, but future conflicts might.
- What might be a cyber equivalent for missions flown, combat missions, and flying hours? These are all criteria Airmen use to understand the experiences of other Airmen. Defensive operators routinely block major attacks and respond to the adversary's changing tactics. Offensive operators intrude into adversary's systems. Each of these can be measured and rewarded and may have an equivalent in cyberspace, which can build cyber heritage and esprit de corps.

- What might be a cyber equivalent for aerial victories and qualification for becoming an ace? Cyber operators, both offensive and defensive, are in routine contact with adversaries looking to do America harm. Sometimes Air Force operators win and sometimes they lose, but the best among them win more consistently. A definition on what constitutes a victory, a concept which is sure to be very elusive, would be one way to celebrate the best traditions of Airmen everywhere.

Nearly 90 years ago, Maj Horace M. Hickam told a doubtful Morrow Board, “I am confident that no general thinks he can command the Navy, or no admiral thinks he can operate an army, but some of them think they can operate an air force.”¹⁸ Today, Airmen are sure they can operate a cyber force but have largely ignored the lessons from the history of cyber conflict and the service’s own cyber heritage. The Air Force must start to inculcate cyber mindedness rooted in history and heritage.

The longer we think cyber conflict is new, the more we will repeat the same mistakes and relearn old lessons. Today’s AF officers learn the Fokker scourge, daylight precision bombing, MiG Alley, and Rolling Thunder. So, must the new Air Force cyber cadre study *yesterday’s* cyber operations to understand those of *tomorrow*? The call to today’s Airmen, and especially the cyber cadre should be clear. Learn your history—know the units, understand the operations, and emulate the aces. And above all, incorporate the lessons. The Air Force used to know this and more. Once it reclaims this heritage, it can lead the world as the premiere force to fly, fight, and win in cyberspace.

Jason Healey

*Director of Cyber Statecraft Initiative
Atlantic Council, Washington, DC*

Notes

1. “Heritage,” US Air Force official website, <http://www.af.mil/information/heritage/index.asp>.
2. Searches conducted on <http://www.airforcehistory.af.mil/main/welcome.asp> and <http://www.afhra.af.mil/>.
3. Gen Ronald R. Fogleman, USAF chief of staff, and Secretary of the Air Force Sheila E. Widnall, “Foreword, Cornerstones of Information Warfare,” *C4I.org*, 1995, <http://www.c4i.org/cornerstones.html>.
4. Maj Gen John P. Casciano, comments to Air Force Association Symposia, 18 October 1996, <http://www.afa.org/aef/pub/la9.asp>.

5. The Air Force had created other cyber units—and was the first service to do so—such as the AF Computer Emergency Response Team and AF Information Warfare Center (AFIWC) in 1993. These critical units, however, did not directly support the war fighter in such a direct way with both offense and defense capabilities. Quote from Atlantic Council event convened by the author on 5 March 2012, “Lessons from Our Cyber Past: The First Military Cyber Units,” <http://www.acus.org/event/lessons-our-cyber-past-first-military-cyber-units>.
6. Quote from Atlantic Council event convened by the author on 5 March 2012, “Lessons from Our Cyber Past: The First Military Cyber Units,” <http://www.acus.org/event/lessons-our-cyber-past-first-military-cyber-units>.
7. Ibid.
8. Other important Air Force heritage that might have been include the first major cyber organization (AFIWC in 1993), that the first AFFOR cyber component was established in 1998 with Col Jim Massaro as AFFOR commander, and that the first real cyber general—that was in cyber jobs from his earliest days as a captain—is the current ACC/A-2, Brig Gen Bradford J. “BJ” Shwedo.
9. Lt Col Dusty Rhoads, “609 IWS: A Brief History, Oct 1995–Jun 1999,” 1.
10. Maj Gen Richard Webber, Comments at 2009 Air Force National Symposium, <http://www.afa.org/events/natsymp/2009/scripts/091119-Webber.pdf>.
11. This is most likely to change as nations put online more physical infrastructure, such as the Smart Grid.
12. Col Ed Crowder, USAF, “Pointblank: A Study in Strategic and National Security Decision Making,” *Airpower Journal*, Spring 1992, <http://www.airpower.maxwell.af.mil/airchronicles/apj/apj92/spr92/crowd.htm>, using information from the AWPDP-1 plan for air war against Germany.
13. Gen William L. Shelton, remarks at Air Force Association, CyberFutures Conference and Technology Exposition, 22 March 2012, audio at <http://www.afa.org/events/CyberFutures/2012/postCyber/default.asp> (quote around 14:47).
14. See Greg Rattray, *Strategic Warfare in Cyberspace* (Boston: MIT Press, 2001).
15. Maj Gen John Casciano, comments at 1996 AFA Symposium.
16. Lt Gen Bob Elder, e-mail to the author, 24 May 2012.
17. Maj Gen John Casciano, e-mail to the author, 11 July 2012.
18. J. S. Shiner, *Foulois and the U.S. Army Air Corps: 1931–1935* (Washington: Office of Air Force History, 1983), 29.

Depleted Trust in the Cyber Commons

Roger Hurwitz

Policymakers increasingly recognize the need for agreements to regulate cyber behaviors at the international level. In 2010, the United Nations Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security recommended “dialogue among States to discuss norms pertaining to State use of ICTs [information and communications technology], to reduce collective risk and protect critical national and international infrastructure.”¹ Since then, the United States, Russia, China, and several other cyber powers have proposed norms for discussion, and in November 2011, the United Kingdom convened an intergovernmental conference to discuss cyber “rules of the road.”² These activities are a positive change from the first decade of this century, when the United States and Russia could not agree on what should be discussed and the one existing international agreement for cyberspace—the Budapest Convention on Cybercrime—gained little traction. Nevertheless, the search for agreement has a long way to go. Homeland Security secretary Janet Napolitano noted in summer 2011 that efforts for “a comprehensive international framework” to govern cyber behaviors are still at “a nascent stage.”³ That search may well be disappointing. Council on Foreign Relations fellows Adam Segal and Matthew Waxman caution that “the idea of ultimately negotiating a worldwide, comprehensive cybersecurity treaty is a pipe dream.” In their views, differences in ideologies and strategic priorities will keep the United States, Russia, and China from reaching meaningful agreements: “With the United States and European democracies at one end and China

Roger Hurwitz, PhD, is a research scientist at MIT’s Computer Science and Artificial Intelligence Laboratory (CSAIL), a senior fellow at the Canada Centre for Global Security Studies at the University of Toronto, and a founder of Explorations in Cyber International Relations (ECIR), a Minerva Research Initiative program at Harvard and MIT. His current work includes the investigation of international cyber norms, the development of computational systems for cyber events data and ontologies, and modeling the complexities of high-profile cyber incidents.

Dr. Hurwitz’s work is funded by the Office of Naval Research. Any opinions, findings, and conclusions or recommendations expressed herein are those of the author and do not necessarily reflect the views of the Office of Naval Research.

and Russia at another, states disagree sharply over such issues as whether international laws of war and self-defense should apply to cyber attacks, the right to block information from citizens, and the roles that private or quasi-private actors should play in Internet governance.”⁴

This essay joins that pessimism on the basis of a more extensive model of the emerging crisis in cyberspace. The essential argument is that maintaining a secure cyberspace amounts to sustaining a commons which benefits all users, but its overexploitation by individual users results in the well-known “tragedy of the commons.”⁵ Here the depletable common resource is trust, while the users are nations, organizations, and individuals whose behaviors in cyberspace are not subject to a central authority. Their actions, which harm the well-being of other users, diminish trust and amount to overexploitation of a common resource. The tragedy of the commons is used repeatedly as an argument for privatization and in retrospect to justify the enclosure movement by English agricultural capitalists in the seventeenth and eighteenth centuries. However, such a tragedy is not inevitable, even when users of a commons are assumed rational in the sense of maximizing self-interest. The late political scientist Elinor Ostrom received the Nobel Prize in economics for determining cases and conditions where, in the absence of government control, users successfully self-organized for sustainable use of a commons.⁶ Unfortunately, as argued below, the current state of cyberspace and its users does not meet most conditions that encourage such self-organization. Both the affordances of the cyber technologies—that is, the way the technologies enable their use—and the mentalities of the users contribute to the unfavorable result.

Embedding the obstacles to international agreements within this wider perspective will highlight the challenging multilayered, complex, and transformative processes that cyberspace presents to states and other entities that would manage it. It is not a passive domain where states can pursue preexisting competitive or conflicting interests, but one whose rapidly changing technologies and applications create opportunities for conflict. It also reasons for cooperation. Accordingly, the next section develops the model of cyberspace as a social system based on a commons—a “socio-ecological system” (SES) and a “common pool resource” (CPR) to use Ostrom’s terminology—that can be sustained but also depleted. The identification of trust as this “resource” and the implications of its depletion will receive particular attention. The third section reviews the variables which Ostrom and her associates have found to encourage self-organization and

evaluates them with regard to cyberspace. The last section considers which of the model variables that currently discourage self-organization could be changed in a more encouraging direction through feasible actions by agents, thus removing some obstacles to reaching international agreements. It also considers how states, absent these changes, might unilaterally respond to cybersecurity crises.

Challenges of the Cyber Commons

Governing a commonly accessible resource, or CPR, is a collective action problem, whether the goal is sustainable exploitation of a fishery or the secure, beneficial use of cyberspace. For natural CPRs, where regeneration of the stock occurs, some limits on individuals' use by amount or kind are needed, lest aggregate use exceed the "carrying capacity." This depletes the resource below the level at which natural processes can sustain it for profitable exploitation. As discussed below, this need for limiting exploitation can also hold for man-made or artificial resources like cyberspace. Limiting or regulating use usually requires a preexisting state or other authority with coercive power, in whose territory the CPR is found—with good reasons. Although the users might recognize the need for limits, individual users are tempted to exceed them in the belief that the added strain on the resource is negligible with regard to its sustainability. Also, individuals who notice their neighbors' violations might be unwilling to punish them for fear of retribution. Nevertheless, Ostrom found many cases where people successfully managed a CPR without the need for state intervention or privatization. In analyzing these, she conceptualizes the CPR as existing within a context of its users' socioeconomic and cultural practices. These practices affect both individual users' choices about exploiting the CPR and the possibility of their collective regulation to sustain it. The CPR and the social context taken together constitute the socioecological system.

One might wonder how a domain can be a commons when every bit of its physical substrate is owned by some organization or a state in contrast, say, to oceans, international airspace, and outer space. Several answers are useful to refining our notion of a cyber commons and any international agreements that would protect it. Lawrence Lessig referred to a model of Internet communication transport that includes layers for the physical substrate, the electronic packets or envelopes for the information, and the

information content itself. He identified the commons with the packet layer, which everyone has a right to access and to which everyone can contribute, so any blocks to the free flow of packets closes the commons.⁷ On this view, the cyber commons is similar to the oceans or international airspace, with its users' primary concern being right of passage.⁸ Lessig and others ultimately grounded this idea of the cyber commons in the human right to access information and express one's opinion. It also resonated with notions of freedom of mobility, global innovation for the Internet, and an evolving worldwide information sphere in which everyone could participate—with the resonance captured in a word: “open.” Endeavors like Wikipedia, the Creative Commons, MIT's free courseware, and the emergent blogosphere could create a second commons—one of content. At the turn of the millennium, Lessig saw such efforts threatened by media content companies, with their broad interpretations of copyright at the expense of fair use and their enlistment of state authorities for draconian treatment of alleged copyright violations. He discounted the argument for a need to protect the intellectual resources from depletion by invoking Thomas Jefferson's image of the candle whose light is undiminished in lighting another candle—a trope for the Enlightenment that encapsulates the promise of the Internet. The unfolding drama was rather that of greedy organizations using the possible misdeeds of a few individuals as a pretext to privatize common intellectual property and undermine the access needed to sustain an Internet culture.⁹

This idea of a “cyber commons” appeared more than a decade ago, when the online population was a tenth of its present size and concentrated in North America and Western Europe, where the Internet was easily seen as another venue in an already rich, lightly regulated, information and communication ecology. It ignored, however, that the Internet was already used by groups in violent struggle against some states—Chechen separatists against Russia—and even liberal states were already proscribing access and distribution of certain information, such as child pornography. Since then, the use of cyberspace, now spilled well beyond the Internet, has become so ubiquitous a national security issue (“securitization”) or a threat to regime stability, that many governments now filter or block certain packet flows, thus replacing the primary cyber commons with their own “safe” enclosures.¹⁰ Nevertheless, the vision of a cyber commons informs significant parts of the cyber policies of the United States and many of its allies and the positions they take with regard to international regula-

tion of cyberspace. Most notable is the State Department's embrace of Internet freedom—the rights of cyber enablement of civic activism—but also significant is the emphasis on global interoperability, noninterference by states with packets passing through their territories, and decisions on Internet technology being made by technologists rather than by political authorities.¹¹

A more identifiable CPR, in keeping with the Ostrom SES model, however, is bandwidth, which can be depleted by spam—an overexploitation of the resource—resulting in degraded delivery of more-valued communications. Spammers have been compared to industrial polluters of natural resource commons because they also pass along to a general public the negative externalities of their actions, whether in the form of users' wait times in a saturated network or added costs for more bandwidth, spam filters, and so forth.¹² The spam phenomenon can be generalized to the consequences of depletion in the general public's "sense of security"; as a by-product of online scams and identity thefts at the individual level; industrial espionage at the organizational level; and infrastructure attacks, like Stuxnet, at the national level. These spur broad demands for cyber-security measures, which are expenses. The provision of these measures, which usually have little effect in stemming the threats, decreases the economic efficiency of cyber-based communications and control. Since the Internet's capability of lowering transaction costs is considered one of its primary benefits for economic and social development, the possible high costs of cyber security are challenging for many states and organizations, perhaps as challenging as the consequences of attacks in the absence of adequate security.¹³

Cyberspace as a Social System

Closely associated with such insecurity is the decline in public or social trust, which might be identified as the ultimate common pool resource in the cyber SES. Jacques Bus follows sociologist Nicolas Luhmann in explaining trust as "a mechanism that reduces complexity and enables people to cope with the high levels of uncertainty and complexity of (contemporary) life." He adds,

Trust expands people's capacity to relate successfully to a real world whose complexity and unpredictability is far greater than we are capable of taking in. In this sense, it is a necessary mechanism for people to live their lives: to communicate,

cooperate, do economic transactions, etc. It enriches the individual's life by encouraging activity, boldness, adventure and creativity, and by enriching the scope of the individual's relationships with others.¹⁴

The notion of public trust, as used here, also includes people's confidence in the institutions, laws, government, and infrastructures of their societies. Public trust with regard to cyberspace encourages individuals and organizations to access and be accessed by one another online, and that in turn enables the network effect in cyberspace; that is, the positive externalities created as more people participate in the network and more interactions occur. This is consistent with findings by social scientists of strong positive correlations between public trust and economic growth.¹⁵

Public trust in cyberspace involves both confidence in the people and organizations individuals deal with through the digital technologies and the trustworthiness of the technologies themselves. Confidence in others online is problematic because those others might be anonymous or only partly identified, and the context of interactions with them is opaque or confusing. It can be buttressed by assumptions about others' concerns for reputation and commitments to roles and by online mechanisms, like certificates and ratings, which can confirm claims made by others. Of late, however, trust in cyberspace may be strained by the publicity for the various cyber threats noted above, organizations' and governments' failures in deterring them, and the compromise of online security mechanisms, like stolen certificates. In addition, public trust suffers from many users' awareness that their online activities are being monitored, whether for commercial exploitation in the West or identification of political dissidents in authoritarian countries.

These abuses may lower or deplete public trust—that is, the aggregate willingness of users to go online—much like overexploitation by some of its users depletes a CPR. On this view, public trust is a rival good whose consumption by a user decreases the amount available for consumption by others. By analogy, continuing abuses against a diminishing public trust could lead to unsatisfactory provision of the online benefits which public trust enables. In concrete terms, individuals and organizations fearing cyber crime, invasions of privacy, and so forth would greatly decrease their use of digital networks for economic transactions, information exchanges, and social interactions. But unlike the usual commons resources, such as forests and fisheries, public trust in cyberspace is not always a rival good. Mutually beneficial online interactions will sustain and increase,

and these are so plentiful at the individual and organizational levels that the abuses are often ignored or quickly forgotten. Consequently, there is little evidence of people exiting cyberspace or avoiding popular sites with controversial privacy policies. Still, in some democratic countries, relevant publics have demanded that service and search providers restrain tracking; some governments have already responded with regulatory policies, which will force adjustments by data aggregators and analysts. These actions can be read as instances of users defending a CPR by turning to existing authority for leadership and norm setting. They show that in addition to security technologies, sustaining trust in cyberspace requires rules, transparent practices, accountability standards, and means of redress acceptable to users. International efforts for agreements to protect and sustain cyberspace will therefore need to take such concerns into account, to some degree. That might not be a formidable challenge. Because cyber “apps” have become indispensable for so many users, they are likely to be reassured, at least momentarily, by small, facile steps by providers or regulators, including policy announcements, opt-out buttons, and new, if unintelligible, service agreements. Put another way, cyberspace is no longer a domain apart from its users, a place to visit at one’s choosing, like a tourist resort, but has penetrated and rewoven the fabric of our lives.¹⁶

Arguably, the spammers, hackers, data collectors, criminal gangs, cyber activists, and state agencies which threaten public trust are not seeking to destroy the Internet or freeze cyberspace—no more than peasants who allegedly overgrazed the commons wanted to degrade it. Ostrom’s work implies two types of agents damage the CPR: poachers from outside the group that maintains the SES and members of the group who exceed their rights to the CPR. By this reckoning, the spammers, cyber criminals, terrorists, and certain activists—for example Lulzsec—would be the poachers in cyberspace. In popular imagination, and sometimes in their own imaginations, they fill the traditional image of pirates—individuals and groups outside nations and beyond the laws of nations.¹⁷ Indeed, some analysts believe that international cooperation to suppress such groups can be easily realized and comprise a first step toward more comprehensive agreements on cyberspace. Of course, as poachers or parasites, these groups are not seeking the demise of cyberspace, since that would put them “out of work.”

The second type includes governments, online service providers, multinational corporations, and others—the so-called stakeholders—who recog-

nize the need for limits but will frequently flaunt such limits in the pursuit of individual interests. Even states that develop cyber weapons to damage cyber-based infrastructures and governments that spy on their online citizens value their own use of cyberspace while planning to constrain its use by others. The resulting ambivalence of many governments is perhaps best captured in a recent Chinese white paper, which celebrates the Internet for enabling economic and social development, notes its use in propagandizing the public and in campaigns against provincial corruption, but stipulates that

no organization or individual may produce, duplicate, announce or disseminate information [on the Internet] having the following contents: being against the cardinal principles set forth in the Constitution; endangering state security, divulging state secrets, subverting state power and jeopardizing national unification; damaging state honor and interests; instigating ethnic hatred or discrimination and jeopardizing ethnic unity; jeopardizing state religious policy, propagating heretical or superstitious ideas; spreading rumors, disrupting social order and stability; disseminating obscenity, pornography, gambling, violence, brutality and terror or abetting crime; humiliating or slandering others, trespassing on the lawful rights and interests of others; and other contents forbidden by laws and administrative regulations.¹⁸

On this view, the strategic problem with the Internet is not its dual use but its many uses. So many, in fact, that unilateral efforts like deep packet inspections to contain the “unwanted uses” themselves threaten the stability and sustainability of cyberspace.

Sophisticated actors who threaten public trust in cyberspace might foresee the adverse consequences of their acts. They might also calculate that whatever the damage they do, the depletion of public trust will be modest or the gains in using the Internet still so great that public trust and mutual accessibility will remain above some minimum threshold. As noted, recent trends support that calculation. Yet, to the point that their conduct cannot be generalized or continue indefinitely—without devastating consequences, that is—to the question, “What if everyone always acted like you?” they must still answer, like Yossarian, “I would be a damned fool not to.” The alternative is for all the Yossarians to act together to change the situation. Is that possible in cyberspace under current conditions? Can a significant number of relevant actors abandon practices that threaten it and commit to rules that sustain it?

Self-Organizing Variables

Ostrom and her associates have identified 10 variables critical for self-organization in a socioecological system—that is, effective and enforced rules of use for a common pool resource in the absence of state authority.¹⁹ Each variable is explained below, sometimes introduced with direct quotations from Ostrom (either italicized or in quotation marks), while manifestation in cyberspace is described and evaluated with regard to its effect on self-organization. Encouraging, discouraging, and neutral effects are indicated by +, −, or 0, respectively. The variables concern properties of the resources being exploited in the SES and characteristics of the user population. In keeping with the observation that public trust in cyberspace depends on the trustworthiness of its hardware and software, as well as the behavior of their users, their properties are considered in evaluating the relevant variables.

As will be seen, Ostrom's explanations of the variables' effects on the possibility for self-organization are consistent with a rational actor model: the probability of self-organization increases the more its contribution to sustaining the common resource exceeds the costs of bringing agents to agreements and enforcing those agreements. Hence, the lower these costs, the greater the probability of self-organization. The assumption with regard to its process is that states through multilateral agreements would set rules and regulations for cyberspace; they would either enforce these directly or empower an international agency to do so.

Size of Resource (−)

Large resources with ill-defined boundaries discourage self-organization because of the high costs of defining the boundaries, monitoring use, and tracing the consequences of malfeasance.

The size of cyberspace, as measured by the several billion devices connected to the Internet, discourages defining its boundaries and monitoring behaviors in it. As a thought experiment, suppose “boundaries” for a trustworthy cyberspace were defined by a centrally maintained giant list of several billion verified safe devices, with “safe” designating malware-free or not having been involved in spying or other penetration operations. This list would require continual updating to accommodate devices being added to the Internet and recurrent verification of the safe devices, because anyone could be vulnerable to attack from a host spoofing a safe device. This approach would be very expensive and only partly effective in

inspiring users' trust; some attacks are so stealthy as to be discovered only well after they have occurred, if at all.

Mapping boundaries and monitoring behavior can be more feasible, affordable, and convincing if national governments assume responsibility for the devices and users in their territories by certifying the machines and credentialing the users. Unilateral and multilateral means could then protect the defined national cyberspaces. Such means include implementations of "national firewalls" and the reduction of national portals, cyber passports for users, and assignment of consecutive IP addresses to specific territories. These steps would not stop all external attacks and exploits within a national cyberspace, but they would facilitate determining the origin of attacks and holding responsible authorities in the state where an attack originated.²⁰

The resulting system would extend the principle of national sovereignty—the cornerstone of contemporary international relations—into cyberspace²¹ and increase a state's control over its residents' online activities. Some states, including a few liberal democracies in the West, have already adopted or advocated some of these measures to deal with cyber security threats. However, many governments, organizations, and individual users will oppose full-blown development of the system for several reasons. First, it would sanction the fragmentation of the Internet into many an "internet in one country" with an attendant constriction of global communications. That process, already foreshadowed in China, Iran, and other authoritarian countries, would set back efforts to build a commons for discussion of items like climate change, scientific knowledge, and medical research on a global agenda. Second, multinational corporations and other agents of globalization, including economic managers in authoritarian countries, will consider this system an obstacle to a global economy in which businesses anywhere can have suppliers and customers everywhere. For them, a particularly threatening aspect of the projection of national sovereignty into cyberspace is the potential restriction in movement of information resources. Third, human rights advocates will oppose conceding the right to define a cyber attack to national governments, since their definitions can include a broad swath of content, as noted above in regard to China, as well as malicious code. Fourth, policymakers are likely to doubt whether governments will accept responsibility for cyber attacks originating in their territories under this system. These doubts can be grounded in

current practices of government claiming ignorance of the attack origins or that they do not have the means to suppress all of them.

Finally, national boundaries in cyberspace are a way of dissecting the commons and privatizing the pieces. Because this commons is a network, its dismantling involves a loss of value. That is, the sum of the values of the parts will be less than the value of the original whole. The loss will be defined in different ways, but its anticipation will motivate broad resistance to the idea of national cyber borders. Nevertheless, the idea brings into relief questions about the character of the cyber commons: whether it is a thin communications overlay on, and ultimately reduced to, diverse geophysical entities and jurisdictions, or does it provide sets of experiences—a mode of being—in which users might acquire new identities transcending national identity. Jacques Bus considers the question, thankfully free of the usual panegyrics about the Internet flattening the world:

Globalization, driven clearly by new ICTs and the Web, creates understanding hence more trust through spreading information on history and reputation of societies, characteristics of societies and the lives of persons living in certain societies, and allowing easy worldwide communication. This may indeed lead to further erosion of the concept of “the human animal is best off at home.” It may well lead to the need for a completely new view on societies and their cohesion and the role trust must play in this.²²

Number of Users (–)

The more users of a CPR, the greater the transaction costs of getting them together and agreeing to change. So group size discourages self-organization, but “its effect on self-organization depends on other SES variables and the types of management tasks envisioned.”

The two billion people who already access the Internet constitute the largest users group in human history. They should have opportunity to express their concerns in any international negotiations on the uses of cyberspace, since in many cases these are likely to be different from those of governments and other powerful stakeholders. For example, users in struggles against their own governments would certainly reject those governments’ representation of their interests regarding anonymity, online tracking, and permitted content. On the other hand, recent world meetings on climate change and on cyberspace itself have demonstrated that processes which are open to groups claiming to represent individual citizens’ interests can rapidly become unmanageable, time consuming, and unproductive. For that reason, an interpretation of national sovereignty,

per which states rightfully represent their citizens' interests, is expedient if not just.

Unfortunately, even this stratagem will not reduce the relevant stakeholders to a manageable number. Negotiations will need to include representation of industrial sectors, especially ICT, and international organizations represented, as well as the states, since these can provide the technical knowledge to inform proposals but can also block implementations of any agreements reached without them. As Ostrom suggests, the number of parties involved might not itself determine the difficulty in reaching an agreement. Rather when more parties are involved, especially when the issues are complex, there will be a greater number of competing claims that take time to reconcile, if they can be reconciled at all. Negotiations for the UN Convention on the Law of the Sea (UNCLOS), which regulates another commons, lasted a decade despite building on centuries of admiralty law and being more confined to issues of state sovereignty. There is much less legal tradition for cyber and, so far, no concerted efforts to harmonize state-level cyber laws. Thus, the very limited and regionally oriented Budapest Convention on Cybercrime has been slow in gaining adherence, with many of its signatories listing numerous reservations.²³ Perhaps some relief from these bleak prospects might be provided by cyberspace itself, in that aggregation of opinions, consultations, and negotiations can themselves now be conducted virtually as well as in person. By organizing information, lowering transaction costs, and speeding communications, cyber tools might permit decision making about their own futures.

Resource Unit Mobility (–)

Due to the costs of observing and managing a system, self-organization is less likely with mobile resource units . . . than with stationary units, such as trees and plants or water in a lake.

Three types of mobility of devices make their effective, actionable monitoring difficult and costly. First, as already noted, the status of a device can change rapidly from “safe” to “compromised,” frequently without the change being discovered until later, if at all. Second, over their course, wide-scale cyber attacks and exploitations will typically deploy different machines located at different IP addresses and geophysical locations. For example, during the massive July 2009 distributed denial of service (DDoS) attack on US government sites, the command and control (C2)

sites reportedly migrated from computers in South Korea to some in Chicago and Berlin. Therefore, any monitoring or defense specific to an attack, like blockading potential C2 sites, will probably involve multiple jurisdictions with consequent problems of coordination. Later investigations will be similarly complicated and attribution inevitably uncertain. As a result, parties to an agreement barring such attacks cannot rely on monitoring to verify that they are complying with the agreement or to identify violators. Third, the rise of mobile computing in the form of laptops, smart phones, and tablets has greatly increased the attack surface of cyberspace and the chore of any future monitoring program. The physical mobility of these devices also means they are exposed over their lifetimes to a variety of cyber threats and surveillance environments and to changes in their own security status. They will be more vulnerable than a machine tethered to a single server within an organization setting that has competent cyber security. They are more liable to penetration, theft of their information, and compromise. Once compromised, they can be turned into carriers for compromising networks to which they later connect, like corporate intranets.²⁴

Importance of Resource to Users (+)

In successful cases of self-organization, users are either dependent on the [resource] for a substantial part of their livelihoods or attach high value to the sustainability of the resource.

An increasing amount of activity throughout the world involves the creation, collection, packaging, use, and distribution of information. The Internet and other parts of cyberspace are vital to these activities. Various government position papers on cybersecurity are clear in recognizing the economic, social, cultural, and scientific importance of cyberspace. In calling for the “creation of a global culture of cybersecurity,” the UN General Assembly recognized that

the increasing contribution made by networked information technologies to many of the essential functions of daily life, commerce and the provision of goods and services, research, innovation and entrepreneurship, and to the free flow of information among individuals and organizations, Governments, business and civil society.²⁵

Even authoritarian regimes in Iran, Egypt, and elsewhere, which confronted massive protests organized by cyber means, have hesitated shutting

down the Internet in their countries because of their economies' dependence on it.

Governments and diplomats, however, have been less clear in recognizing how foundational public trust is for cyberspace. In calling for discussions of international norms for cyberspace, the UN group of governmental experts took mainly a national security perspective: Cyber crime and other cyber threats are disruptive to government, economic, and social functions; lack of a common understanding of the intents behind certain behaviors in cyberspace can lead to conflicts which might escalate to threaten international security.²⁶

Productivity of System (+)

If [a resource] is already exhausted or very abundant, users will not see a need to manage for the future. Users need to observe some scarcity before they invest in self-organization.

The growth of cyber crime, the incidence of attacks and exploits, the proliferation of malware, and threats to critical cyber infrastructure have raised questions whether the benefits of cyberspace can be sustained under present security practices. These questions clearly motivate the various calls for international agreements on cyberspace behavior. Jacques Bus notes that the possibility of states being behind many cyber threats “proves the urgency to come to international agreements on restraints in and defense against cyber attacks and for international cooperation to bring it under control.”²⁷ Having identified public trust as the depletable resource in cyberspace, Bus continues, “Public and private sector must work together at the international level to build a well balanced infrastructure of technology and law/regulation that will give citizens trust to use the opportunities of the new digital world.”²⁸ In a speech to the 2011 Munich Security Conference, British foreign minister William Hague made similar connections:

We are working with the private sector, to ensure secure and resilient critical infrastructure and the strong skills base needed to seize the economic opportunities of cyber space, and to raise awareness of online threats among members of the public. But being global, cyber threats also call for a collective response. In Britain we believe that the time has come to start seeking international agreement about norms in cyberspace.²⁹

Predictability of System Dynamics (0)

System dynamics need to be sufficiently predictable that users can estimate what would happen if they were to establish particular . . . rules or no-entry territories.

The consequences of a continuing lack of international regulation are more predictable than the effect of agreement and monitoring for some standards of behavior. With deterioration of public trust in cyberspace, the expansion of use—in terms of time spent, applications, and dependencies—will decelerate, and that will be accompanied by lower growth or drop in the incentives for development. Some users may have already reduced their use of public networks for critical data transmission; some organizations have reduced the number of access points or portals to themselves. These steps might grow toward widespread delinking and fragmentation—phenomena which devalue cyberspace.

Projecting the loss in value of a vulnerable cyberspace compared to a safe one is problematic because of different models for evaluating the socio-economic value of cyber networks. However, it seems reasonable to suppose that as new users are drawn more from lower economic strata and less-developed countries, the economic value of the networks will increase at a lower rate than in earlier stages of their growth.³⁰ Such a trend has mixed implications for self-organization. First, providers will have little incentive to increase their investments in cyber security—especially if security costs are a linear function of the number of users. But inaction by the providers could put more pressure on governments to work for agreements that reduce threats. On the other hand, the trend also suggests that any exit of users will not initially diminish network value. So, until the situation is deemed intolerable and not just bad, governments, mindful of the costs of agreements, could resist pressure and delay self-organizing, despite their public calls for action.

Leadership (0)

When some users of any type of resource system have entrepreneurial skill and are respected as local leaders as a result of prior organization for other purposes, self-organization is more likely.

Leadership is lacking for potentially productive, state-level negotiations, but not for want of actors that have had roles in organizing cyber-

space. Over the past decade, the Internet Corporation for Assigned Names and Numbers (ICANN) has provided competent, although frequently criticized, administration of domain allocations and oversight of registration. It has accommodated the spectacular growth of the Internet and accompanying commercial demands with a redesign of policies for top-level domains. While it has not been particularly open to the grassroots participation specified in its multistakeholder model, it has retained the confidence of service providers and the respect of most states, as evidenced by the UN's restraint from seeking involvement in administration of the Internet. But the ICANN is no norms entrepreneur and lacks the political skills and leverage to reconcile competing interests among states over cyber behaviors and security. Additionally, it is seen by many states as a tool of US policy.

The Internet Engineering Task Force (IETF) has exercised leadership in Internet protocols, mostly as the endorser of standards. Its own history exemplifies self-organizing among stakeholders for management of a commons, but its amorphous decision-making process is an awkward model for negotiations on constraining human activities. In any case, it is unqualified to lead in such negotiations, its ambit is limited to the technical realm, its centrality in that realm has diminished as concerns now focus more on mobile computing apps and other layers beyond its purview, and its membership is still heavily American and European.³¹

The International Telecommunications Union (ITU), the UN agency responsible for ICT, has the ambition to lead policymaking and administration of cyberspace, and it led in organizing the World Summits on the Information Society (WSIS), which focused on soft issues: development-oriented uses of cyberspace, Internet governance, bridging digital divides. Seen in the West as a tool for Russian and Chinese policy interests, it lacks the political credibility to assume leadership on hard issues like cyber espionage, information rights, and so forth. It probably also lacks the technological competence; the cybersecurity standards it developed and promoted in collaboration with the International Organization for Standardization (ISO) have proved expensive and unworkable.

Norms/Social Capital (+)

If users share norms of reciprocity and sufficiently trust one another to keep agreements, they will face lower transaction costs in reaching agreements and monitoring. Continued economic globalization and the ab-

sence of major interstate wars could suggest that the major powers are developing adequate reciprocity structures and conflict avoidance mechanisms. Indeed, this assessment is supported by the fears expressed in the calls for cyber norms that misunderstandings about cyberspace behaviors could trigger unwanted conflicts. Nevertheless, the failure of negotiations on environmental regulations raises doubts that negotiations over cyberspace can fare any better, especially since the major powers have ideological differences regarding cyberspace, as great as the differences among economic interests that block resolutions of environmental issues.

Broadly speaking, the Russian and Chinese policymakers seek to extend the principle of national sovereignty to cyberspace by establishing a norm of the state being the final arbiter of matters relating to cyberspace in its territory.³² From a Western perspective, their motives are to control the ideational space that cyber networks afford their populations and to prevent inquiry into use of cyber by their governments or proxies for military campaigns, political espionage, industrial espionage, and crime. Recall, however, that the political traditions in Russia and China, even in the pre-Communist days, empowered state authorities to decide what their citizens should think, and that the principle of national sovereignty bars outsiders from interfering with the exercise of that power. Furthermore, Russian officials are keenly aware that Chechen insurgents or terrorists have used cyber technologies in their violent struggles against Russia. So an uncontrolled Internet can be politically threatening and easily exploited by external rivals, in particular the United States. For example, when cyber-fueled protests occurred in Russia, premier, presidential candidate, and target of the protests, Vladimir Putin, branded these protests the work of “foreign enemies.”³³ On this view, outsiders enabling dissent within a country is no contribution to public debate; it is “information warfare” conducted to weaken regimes to the point of greater accommodation with the outsiders or even collapse. Already, in 2008, Russia, China, and other members of the Shanghai Coordination Organization (SCO) have agreed to outlaw supporting or hosting the dissemination of potentially disruptive information. In September 2011, in seeming response to foreign governments’ and Diasporas’ support for cyber activism in the Arab world, Russia proposed that countries log the online activities of their residents suspected of such disseminations.

In contrast, the United States and its NATO allies tend in their pronouncements to view cyberspace as a central institution for a global

economy, a means for worldwide scientific and cultural exchange, a commons for political debate and development, and a social medium. Given this variety of functions, there follows a multistakeholder model for control and defense of cyberspace, with states being one type of stakeholder, along with nongovernmental organizations, service providers, ICT companies, critical infrastructure entities, corporate users, and individual users. But because cyberspace, particularly the Internet, is prey to attacks and exploits by criminals, terrorists, and even states, by virtue of their authority and capabilities, states have primary responsibility to provide the needed security without harming the interests of other stakeholders. The diffusion of norms and treaties, such as the Budapest Convention on Cybercrime, are instruments for fulfilling such responsibility, as are the nurturing of a cyber-security culture and capabilities around the globe.³⁴

This view, wedded to a decade-old vision of the Internet, ignores the demographic and technological changes that are remaking cyberspace and expectations for it: the change from hundreds of millions of users concentrated in North America and Europe connected to the Internet through computers to billions of users with the bulk in south and east Asia connected through mobile devices and the rise of an Internet of things. As a result, practices that might have once seemed in the interest of all are now controversial and contested.³⁵ India, Brazil, and South America—leading voices on cyber issues among “nonaligned” countries—want these changes to be acknowledged as conceded major parts in any negotiations. They consequently favor transfer of authority away from technologically oriented agencies, reflecting the multistakeholder model, including ICANN and IETF, to a more policy-oriented agency, possibly under the UN, though not necessarily the ITU, that gives every state an equal voice.

Knowledge of the Socioeconomic System (+)

When users share common knowledge of relevant SES attributes, how their actions affect each other and rules used in other SESs, they will perceive lower costs of organizing.

The various calls for cyber rules reflect policymakers' knowledge that certain behaviors disrupt normal activities, sow public distrust, and threaten the sustainability of cyberspace. Their willingness to discuss issues beyond cyber crime acknowledges that those misbehaving may include their own governments and citizens. So, less time and money are needed to raise

consciousness or convince skeptics that a problem exists and international cooperation can help solve it. Choosing what to do requires more knowledge of the dependencies among various processes in cyberspace, particularly how the technological affordances affect social (agents') behaviors. The efforts at environmental regulation show that broad, comprehensive solutions will be opposed even when those who feel threatened by the proposal are offered side payments. So the problem space has to be decomposed with selection of some target whose proposed solution could gain traction, help reduce the overall level of cyber insecurity, and build confidence among the various agents, thus enabling pursuit of other targets. One frequent suggestion is that states cooperate to suppress cyber criminal gangs by denying their means to monetize their thefts. This suggestion understands (a) the gangs' dependency on particular banks and (b) that cyber crime serves as a development lab and testing ground for malware that might later be used by intelligence agencies in some states. Less known is how strongly these agencies depend on the gangs and, therefore, the incentives their states need to cooperate on the proposal.

Collective Choice Rules (0)

When users have full autonomy at the collective-choice level to craft and enforce some of their own rules, they have lower transaction costs as well as lower costs in defending the resource against invasion by others.

This variable implies that the more people can see themselves as authors of the rules they are expected to follow, the more they will follow those rules. This result is important for cyber security and public trust in cyberspace, because good "computer hygiene" at the organizational and individual levels can blunt a considerable amount of computer crime and exploits, perhaps as much as 80 percent.³⁶ Unfortunately, the number of users and the diffuseness of their representation would seem to preclude public participation in making rules, as mentioned before. Consequently, users will be less able to see their rule following as part of a global interdependent effort to sustain cyberspace and therefore their own benefit from it. The top-down directives they receive will more likely justify the rules only in terms of protecting the individual or organization.

Changing Variables and Crisis Response

The values of the Ostrom variables, summarized in the table below, do not favor self-organization in the cyber SES. Conditions are not ripe for productive, enforceable agreements under which stakeholders, especially states, limit their trust-eroding cyber behaviors. As indicated by the positive values for the “importance of the resource” and “productivity of the system” variables, the widespread expressions of fear for the future of cyberspace has sparked interest in such agreements. However, nothing beyond that should be expected until the values of some technological and other social variables change. Arguably, the pursuit now of a comprehensive global agreement or fallback to agreements among the “like-minded” will be counterproductive. It will likely deepen distrust among major cyber powers and discourage the sharing of useful knowledge of the cyber SES. That seems to be the primary outcome of the recent London conference on cyber “rules of the road.”³⁷

Variable	Value
Size of resource	—
Number of users	—
Resource unit mobility	—
Importance of resource	+
Productivity of system	+
Predictability of system dynamics	0
Leadership	0
Norms/social capital	+
Knowledge of SES	+
Collective choice rules	0

Several feasible measures could improve prospects for effective agreements and/or sustain public trust in cyberspace. Consider the following changes.

Develop Global Identity Management

Jacques Bus recommends the development of a “globally interoperable trustworthy system for Identification and Authentication” as essential for

trust among Internet users.³⁸ States, including some liberal democracies, are already requiring verified identification from Internet users. Interoperability of local standards would facilitate, if needed, the identification of a user of an Internet-linked device anywhere. Users could retain some anonymity or privacy under this regime, since different sites and transactions would demand different degrees of disclosure. Authoritarian regimes could more easily identify people in cyber networks of resistance, but they might find they are better off not identifying nonviolent resisters, while trying to identify and suppress violent ones. That strategy could channel opponents toward the nonviolent networks and give the regimes more breathing room. Their restraint in this regard could enable states that support their opponents to cooperate in the identification system. In terms of the Ostrom variables, identity management reduces some of the deleterious effects of resource mobility.

Increase Public Participation on Cyber Security

Discussions of cyber security policies in informed, relevant publics can have the double effect of putting pressure on respective national governments and involving these publics in rule-making processes. The UN resolution for the “creation of a global culture of cybersecurity” anticipates that national cyber security efforts will have broad societal involvement, including that of the private sector, civil society, academia, and private individuals, but it is silent regarding rule-making roles for nongovernmental actors. The public-private partnerships that have already emerged in Europe and North America appear focused on coordinating organization-level efforts and sharing information, without critiquing or innovating policies. But nongovernmental members, particularly any transnational corporation (TNC) and international nongovernmental agency (INGO), for example Freedom House, should be encouraged to suggest rules. Many have experienced cyber attacks in a variety of legal and technological environments and probably know better than observers or governments what cyber laws and practices need to be harmonized across countries as part of international agreements.

The Internet Governance Forum (IGF), a consultative body established by the UN and based on a multistakeholder model, might also be used for public input into global-level conversations on rules for cyberspace. Its meetings have discussed cyber security issues but have so far deferred to national governments and specialized agencies for policy proposals. But

the IGF could use cyber tools and techniques, such as online surveying and crowd sourcing to collect and aggregate public opinion about rules and regulations needed in any future agreements.

Confidence Building through International Cooperation on an “Easy” Task

Although comprehensive agreements on cyberspace behaviors might be unattainable, international cooperation on some cyber threats and emergencies can be strong and effective, for example, the worldwide response to the Conficker worm or the working alliance of the Japan, China, and South Korea CERTs. In these cases, the cooperation builds upon “invisible norms” or commitments shared among cyber technologists, but it can give onlooking policymakers some confidence about their countries’ working together on cyber problems. So, their confidence could grow with more cases where a challenge triggers a widely shared professional commitment and the ensuing cooperation achieves some success. Some cyber crimes seem suitable candidates for the challenge, notably child pornography, low-level fraud, and identity theft. There is, however, a need for some agency to take the lead in promoting the urgency of suppressing the chosen crime.

This essay has used economic reductionism to argue that conditions are not ripe for reaching and enforcing international agreements on the uses of cyberspace. The argument holds that if people who exploit a commons know that overexploitation will degrade that commons they can agree to limit their behavior, providing the costs of coming to agreement and enforcing it are affordable. In this argument, self-limitation is in service to self-interest—to sustain one’s benefits from the commons. As far as the actor, whether individual, organization, or nation is concerned, cyberspace is just another domain where it pursues its self-interest. Cyberspace is, of course, much richer. It has become the basis and means for reorganizing much of contemporary social, economic, cultural, and intellectual life in developed countries. It provides a principal means for a global conversation about shared issues. To the extent it retains public trust, cyberspace cultivates new social bonds and identities that augment preexisting ones, like nationality. For all that, it commands some allegiance.

Even its advocates do not think an international cyber treaty would sufficiently protect states, organizations, and individuals from the various attacks arising in cyberspace. Although a treaty would be a restraint on its

signatories and facilitate sanctions of its violators, adequate cyber defense at the state level would still require resistance (hardening) of digital networks, especially those supporting critical infrastructure; resilience of organizations likely to be attacked; and reasonable deterrence with respect to nonsignatories. In the absence of international agreement(s), reliance on these other components would increase moderately. Furthermore, because digital networks are necessary for economic globalization, states will continue to cooperate on the technical plane and with regard to Internet governance at least to the point of assuring interoperability at the global level. Such cooperation will not extend to control industrial espionage, protect critical information infrastructures or assure information freedom, three issues which have recently emerged as foci of distrust among states. These and other cyber issues at the international level will likely be addressed in the midterm future in disjointed and incremental fashion—the strategy of muddling through. These are not necessarily bad results, and few users will experience any loss of benefits from cyberspace. On the other hand, the insecurity there will persist, and the opportunity to build public trust on a global level will have passed. 

Notes

1. UN General Assembly, “Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” A/65/201, 30 July 2010, <http://www.unidir.org/pdf/activites/pdf5-act483.pdf>.

2. For a review of the London conference, see Peter Apps, “Disagreements on Cyber Risk East-West ‘Cold War,’” *Reuters*, 2 February 2012, <http://www.reuters.com/article/2012/02/03/us-technology-cyber-idUSTRE8121ED20120203>.

3. “Remarks by Secretary Napolitano before the Joint Meeting of the OSCE Permanent Council and OSCE Forum for Security Cooperation,” Department of Homeland Security news release, 1 July 2011, <http://www.dhs.gov/ynews/speeches/2011-napolitano-remarks-osce-council-austria.shtm>.

4. Adam Segal and Matthew Waxman, “Why a Cybersecurity Treaty Is a Pipe Dream,” *Council on Foreign Relations*, 27 October 2011, <http://www.cfr.org/cybersecurity/why-cybersecurity-treaty-pipe-dream/p26325>.

5. See G. Hardin, “Tragedy of the Commons,” *Science* 162 (1968): 1243–48, for the classic formulation of the argument.

6. Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge, UK: Cambridge University Press, 1990); and Ostrom et al., “A General Framework for Analyzing Sustainability of Social-Ecological Systems,” *Science* 325, no. 5939 (24 July 2009): 419–22.

7. Lawrence Lessig, “The Public Domain,” *Foreign Policy*, 30 August 2005, http://www.foreignpolicy.com/articles/2005/08/30/the_public_domain.

8. See for such analogy Abraham Denmark and James Mulvenon, eds., *Contested Commons: The Future of American Power in a Multipolar World* (Washington: Center for a New American Security, 2010).

9. In using the society of the English village commons as their governing metaphor, advocates of an Internet where information flows freely may have tended toward an idyllic or prelapsarian vision. In a dismissive review of Lewis Hyde, *Common as Air* (New York: Farrar, Straus, and Giroux, 2010), the work of one such advocate, David Wallace-Wells, quotes E. P. Thompson's assessment in his classic *The Making of the English Working Class* (New York: Vintage Books, 1966) that English agrarian culture before enclosure was "intellectually vacant . . . and plain bloody poor." Ignoring that enclosure forced people off the land and did not improve the lot of those who remained, Wallace-Wells argues by analogy that we are doomed to cultural sterility without the enclosures of broad copyright in "The Pirate's Prophet: On Lewis Hyde," *Nation*, 15 November 2010, <http://www.thenation.com/article/155619/pirates-prophet-lewis-hyde?page=0,0>.

10. For the securitization of cyber in the United States, see M. Dunn Cavelty, *Cyber-Security and Threat Politics: US Efforts to Secure the Information Age* (New York: Routledge, 2008). For types and extent of enclosure practices, see Ronald Deibert et al., eds., *Access Denied: The Practice and Policy of Global Internet Filtering* (Cambridge: MIT, 2008); and Deibert et al., eds., *Access Controlled* (Cambridge: MIT, 2010).

11. *International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World* (Washington: The White House, May 2011), http://www.whitehouse.gov/sites/default/files/rss_viewer/international_strategy_for_cyberspace.pdf. See also Secretary of State Hillary Clinton, "Remarks on Internet Freedom," 21 January 2010, <http://www.state.gov/secretary/rm/2010/01/135519.htm>. The State Department's high-profile decry of foreign governments' politically motivated filtering led it to oppose the congressional antipiracy bills (SOPA and PIPA), which would have mandated commercially motivated filtering of foreign sites.

12. "Jo Twist, Web Guru Fights Info Pollution," *BBC News*, 13 October 2003, <http://news.bbc.co.uk/2/hi/technology/3171376.stm>. Another type of inordinate bandwidth consumption, the distributed denial of service, is intended to directly inflict some types of costs, such as reputational, financial, or political, on its target by forcing the target's web servers to crash under the crush of demands for service. DDoS can rise to the level of a national security matter, as exemplified in the 2007 attack on Estonian government and critical infrastructure websites.

13. Discussion of the obstacles and costs of "adequate" security for the current inherently vulnerable technologies of cyberspace are beyond the present scope. In addition to costs for cyber security personnel, they include much-less-estimable costs for revamping organizational cultures. Many firms, especially in the financial sectors, have reportedly chosen to defer such costs and to treat any loss to cyber crime or espionage as costs of doing business, while making efforts to suppress publicity of such losses for fear of the costs to their reputations. As this article goes to press, I learned that L. Jean Camp, "Reconceptualizing the Role of Security User," *Daedalus* 140, no. 4 (2011): 93–107, also applies Ostrom's analytic of self-organization to the challenge of cyber security. Camp's focus, however, is on the possibilities of individual end users forming small-scale communities in which information sharing on cyber threats and cyber hygiene are effectively practiced.

14. Jacques Bus, "Societal Dependencies and Trust," in Hamadoun Touré et al., *The Quest for Cyber Peace* (Geneva: International Telecommunications Union, 2011), 18.

15. Ibid., 19, citing Francis Fukuyama, *Trust: The Social Virtues and the Creation of Prosperity* (New York: Free Press, 1995), and Robert Putnam et al., *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton, NJ: Princeton University Press, 1993). For a negative example, see Anthony Padgen, "The Destruction of Trust and Its Economic Consequences in the Case

of Eighteenth-Century Naples,” in *Trust: Making and Breaking Cooperative Relations*, ed. Diego Gambetta (London: Basil Blackwell, 1988), 127–41.

16. Iranians’ use of the TOR anonymizing networks suggests that some users need cyber so much that even a small amount of reassurance will induce them to return to using previously compromised applications, despite the risks involved. The graphs for usage are spiked, showing that immediately after Iranian authorities announce a blockade or monitoring of a particular TOR site, the number of Iranian users on the network drops precipitously. It picks up again after TOR developers announce a workaround to the Iranian measures. See <https://metrics.torproject.org/users.html?graph=direct-users&start=2010-11-28&end=2012-02-26&country=ir&dpi=72#direct-users>.

17. Daniel Heller-Roazen, *The Enemy of All: Piracy and the Law of Nations* (Cambridge: MIT Press, 2008).

18. Information Office of the State Council of the People’s Republic of China, “The Internet in China,” 8 June 2010, http://www.china.org.cn/government/whitepaper/node_7093508.htm.

19. Elinor Ostrom, “General Framework for Analyzing Sustainability of Social Ecological Systems,” *Science* 325 (24 July 2009): 419–22.

20. A view of “state responsibility” is elaborated in the Russian draft for a “Convention on International Information Security,” presented to the Second International Meeting of High-Level Officials Responsible for Security Matters, Ekaterinburg, Russia, 22 September 2011, <http://2012.infoforum.ru/2012/files/konvencia-mib-en.doc>. A problem with any plan that assigns responsibility to states for the cyber behaviors of their residents is that many states lack cyber security awareness, capacity, and computer forensic capabilities. This problem and the role for technologically advanced nations to help less-advanced ones build such capacity are recognized in the US International Strategy for Cyberspace and the UN General Assembly Resolution “Creation of a global culture of cybersecurity and taking stock of national efforts to protect critical information infrastructures,” A/Res/64/211, 17 March 2010, <http://www.citizenlab.org/cybern norms/ares64211.pdf>.

21. Chris Demchak and Peter Dombrowski, “Rise of a Cybered Westphalian Age,” *Strategic Studies Quarterly* 5, no. 1 (Spring 2011): 32–61.

22. Bus, “Societal Dependencies and Trust,” 21.

23. Stein Schjøberg, “Wanted: a United Nations Cyberspace Treaty,” in Andrew Nagorski, ed., *Global Cyber Deterrence: Views from China, the U.S., Russia, India, and Norway* (New York: EastWest Institute, 2010), 11.

24. Ellen Nakashima and William Wan, “In China, Business Travelers Take Extreme Precautions to Avoid Cyber-Espionage,” *Washington Post*, 26 September 2011, http://www.washingtonpost.com/world/national-security/in-china-business-travelers-take-extreme-precautions-to-avoid-cyber-espionage/2011/09/20/gIQAM6cR0K_story.html. See also Joel Brenner, *America the Vulnerable* (New York: Penguin Press, 2011), 61ff.

25. UN General Assembly Resolution 64/211: “Creation of a global culture of cybersecurity and taking stock of national efforts to protect critical information infrastructures,” <http://www.citizenlab.org/cybern norms/ares64211.pdf>.

26. UN General Assembly Resolution 65/201: “Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security,” <http://www.unidir.org/pdf/activites/pdf5-act483.pdf>.

27. Touré et al., *Quest for Cyber Peace*, 16.

28. Ibid., 25.

29. Foreign Secretary William Hague, “Security and Freedom in the Cyber Age—Seeking the Rules of the Road,” speech to the Munich Security Conference, 4 February 2011, <http://www.fco.gov.uk/en/news/latest-news/?view=Speech&id=544853682>.

30. According to the well-known Metcalfe’s law, the value of a network is proportional to the number of cross connections among its N users, that is N^2 . The growth (or decline) in value with each user who joins (leaves) the network is proportional to 2^N . The more extreme Leek’s law equates network value with the number of distinct audiences that can be formed from the number of users, i.e., the number of subsets less the null set of N or $2^N - 1$. So the value of the network would incredibly double (or be halved) with each user joining (or leaving). A more reasonable evaluation, especially for large networks, assumes differential use by those in the network. Consistent with power laws (long-tail phenomena), usage is assumed to decline exponentially with delay in joining the network. Usage or transactions over the N users describes a hyperbole, with the first joiners the heaviest users. The cumulative benefit, hence value of the network, is then proportional to the area under the curve or natural log of N ($\ln N$). The increase (decrease) in network value with each person joining (leaving) is significantly less than estimated by Metcalfe’s law, and the change is decreasing rather than increasing. Thus, if the network provider’s cost of acquiring an additional user is fixed, a point of diminishing returns on value will be reached.

31. My thanks to Phillip Hallam-Baker for discussion of this point.

32. Ekaterinburg draft. (see note 20).

33. Michael Bohm, “Putin Chasing Imaginary American Ghosts,” *Moscow Times*, 9 February 2012, [http://www.themoscowtimes.com/opinion/article/putin-chasing-imaginary-ghosts/452802.html](http://www.themoscowtimes.com/opinion/article/putin-chasing-imaginary-american-ghosts/452802.html)<http://www.themoscowtimes.com/opinion/article/putin-chasing-imaginary-american-ghosts/452802.html>.

34. See UN General Assembly Resolution 64/211: “Creation of a global culture of cybersecurity and taking stock of national efforts to protect critical information infrastructures,” adopted 17 March 2010.

35. Ronald Deibert and Rafal Rohozinski, “Contesting Cyberspace and the Coming Crisis of Authority,” in Deibert et al., *Access Contested*, 21–41.

36. Brenner, *America the Vulnerable*, 239–44; and Brenner, personal communication, 2010.

37. Apps, “Disagreements on Cyber Risk East-West ‘Cold War.’”

38. Bus, “Societal Dependencies and Trust,” 24.

Escalation Dynamics and Conflict Termination in Cyberspace

Herbert Lin

US national security planners have become concerned in recent years that this country might become engaged in various kinds of conflict in cyberspace. Such engagement could entail the United States as the target of hostile cyber operations, the initiator of cyber operations against adversaries, or some combination of the two.

To date, most serious analytical work related to cyber conflict focuses primarily on the initial transition from a preconflict environment to that of conflict. Little work has been done on three key issues: (1) how the initial stages of conflict in cyberspace might evolve or escalate (and what might be done to prevent or deter such escalation), (2) how cyber conflict at any given level might be deescalated or terminated (and what might be done to facilitate deescalation or termination), and (3) how cyber conflict might escalate into kinetic conflict (and what might be done to prevent kinetic escalation). Each of these issues is important to policymakers, both in preparing for and managing a crisis. Before beginning that discussion, it is instructive to consider some relevant terminology and concepts.

Terminology and Basic Concepts

The term *offensive cyber operations* as used here refers collectively to actions taken against an adversary's computer systems or networks that harm the adversary's interests. In general, an offensive cyber operation

Dr. Herbert Lin is chief scientist at the Computer Science and Telecommunications Board, National Research Council of the National Academies, where he has been study director of major projects on public policy and information technology. Of particular note is his role as editor of a 2009 NRC study on cyber attack as an instrument of national policy and a 2010 study on cyber deterrence. He previously served as staff scientist for the House Armed Services Committee (1986–90), where his portfolio included defense policy and arms control issues. He received his doctorate in physics from MIT.

This article is largely based on chapter 9 of *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities* by William Owens, Kenneth Dam, and Herbert Lin (Washington: National Academies Press, 2009). The author is solely responsible for any deviation from the conclusions of that report. This article does not necessarily reflect the views of the research sponsors, the MacArthur Foundation or the Microsoft Corporation.

gains access to an adversary's computer system or network and takes advantage of a vulnerability in that system or network to deliver a payload. In a non-cyber analogy, *access* might be any available path for reaching a file in a file cabinet. A *vulnerability* might be an easy-to-pick lock on the file cabinet—and note that ease of picking the lock is irrelevant to an Earth-bound intruder if the file cabinet is located on the International Space Station where access to the file cabinet would be difficult. The *payload* describes what is to be done once the intruder has picked the lock. For example, the intruder can destroy the papers inside, alter some of the information on those papers, or change the signature on selected documents.

Access is “easy” when a path to the target can be found without much difficulty; a computer connected to the Internet may well be such a target. Access is “difficult” when finding a path to the target is possible only at great effort or may not be possible for any practical purposes. An example of such a target may be the onboard avionics of an enemy fighter plane, which is not likely to be connected to the Internet for the foreseeable future. In general, access to an adversary's important and sensitive computer systems or networks should be expected to be difficult. Furthermore, access paths to a target may be intermittent—a submarine's on-board administrative local area network would necessarily be disconnected from the Internet while underwater at sea but might be connected while in port. If the administrative network is ever connected to the on-board operational network (controlling weapons and propulsion) at sea, an effective access path may be present for an adversary.

A *vulnerability* is a security weakness in the system or network that is introduced by accident (by some party that has a legitimate reason to access the system) or on purpose (by a would-be intruder). An accidentally introduced weakness (a “security bug”) may open the door for opportunistic use of the vulnerability by an adversary. Many vulnerabilities are widely publicized after they are discovered and may be used by anyone with moderate technical skills until a patch can be disseminated and installed.¹ Adversaries with the time and resources may also discover unintentional defects that they protect as valuable secrets—also known as *zero-day vulnerability*.² A deliberately introduced vulnerability occurs because the intruder takes an action to create one where one did not previously exist. For example, an intruder might deceive a legitimate user of the targeted system or network to disable a security feature (e.g., reveal a password). Both kinds of vul-

nerability are useful to intruders as long as the weaknesses introduced remain unaddressed.

Payload is the term used to describe the things that can be done once a vulnerability has been exploited. For example, once a software agent (such as a virus) has entered a given computer, it can be programmed to do many things—reproduce and retransmit itself, destroy files on the system, or alter files. Payloads can have multiple capabilities when inserted into an adversary system or network—that is, they can be programmed to do more than one thing. The timing of these actions can also be varied.

Depending on the intent of the intruder, an offensive cyber operation can be classified as cyber attack or cyber exploitation. *Cyber attack* is the use of deliberate information technology (IT)-related actions—perhaps over an extended period of time—to alter, disrupt, deceive, degrade, or destroy adversary computer systems or networks or the data and/or programs resident in or transiting these systems or networks.³ Such effects on adversary systems and networks may also have indirect effects on entities coupled to or reliant on them. A cyber attack seeks to cause adversary computer systems and networks to be unavailable or untrustworthy and therefore less useful to the adversary. Because so many different kinds of cyber attack are possible, the term *cyber attack* should be understood as a statement about a methodology for action—and that alone—rather than as a statement about the scale of the effect of that action. *Cyber exploitation* is the use of deliberate IT-related actions—perhaps over an extended period of time—to support the goals and missions of the party conducting the exploitation, usually for the purpose of obtaining information resident on or transiting through an adversary's computer system or network. Cyber exploitations do not seek to disturb the normal functioning of a computer system or network from the user's point of view—indeed, the best cyber exploitation is one that goes undetected.

The similarity between these two concepts and the exploitation channel are the most important characteristics of offensive cyber operations. *Cyber attack* and *cyber exploitation* are very similar from a technical point of view. They use the same access paths and take advantage of the same vulnerabilities; the only difference is the payload they carry. These similarities often mean that the targeted party may not be able to distinguish easily between cyber exploitation and cyber attack—a fact that may result in that party's making incorrect or misinformed decisions. The primary technical requirement of cyber exploitation is that delivery and execution of its payload be

accomplished quietly and undetectably. Secrecy is often far less important when cyber attack is the mission, because in many cases the effects of the attack will be immediately apparent to the target. All exploitation operations require a channel for reporting the information they collect. If the channel happens to be two-way, payloads can be remotely updated. Thus, the functionality of the operation may be different today than it was yesterday—most significantly, it may be an exploitation payload today and an attack payload tomorrow. In some cases, the initial payload consists of nothing more than a mechanism for scanning the system to determine its technical characteristics and an update mechanism to retrieve the best packages to further the compromise.

Attribution

Attribution is the task of identifying the party that should be held politically responsible for an offensive cyber operation.⁴ *Technical attribution* is the ability to associate an attack with a responsible party through technical means based on information made available by the cyber operation itself—that is, technical attribution is based on clues available at the scene (or scenes) of the operation. *All-source attribution* is a process that integrates information from all sources, not just technical sources at the scene of the attack, to arrive at a judgment (rather than a definitive and certain proof) concerning the identity of the intruder.

As a general rule, attribution is a difficult matter. It becomes more difficult as more of the following factors are present:

- The techniques used have never been seen before, so the investigator is unable to link them to other parties that have used similar techniques in the past.
- The intruder leaves no forensic clues and makes no technical mistakes (i.e., tradecraft is error free).
- The intruder maintains perfect operational security, so there are no other sources of intelligence (e.g., SIGINT, HUMINT).
- The motivations for conducting the operation are unknown, or the operation occurs during a time when political circumstances do not suggest conflict or adversarial relations to associate a known party's demands or interests with a possible perpetrator.

- The intrusion requires a rapid response which prevents a thorough investigation, raising the likelihood of a mistaken attribution.

If most or all of these factors are present, then attribution is virtually impossible. On the other hand, it is rare that *all* of these factors are present. One might thus reasonably conclude that although technical attribution is indeed difficult, all-source attribution is sometimes possible. Solving the problem of attribution is not as hopeless as is often portrayed.

The Need for Intelligence Support

Offensive cyber operations against a given system require detailed knowledge about both access paths to and vulnerabilities in the targeted system. The amount of detail should not be underestimated—in principle, it may involve very “small” details such as

- the specific processor model (and even the serial number of the processor) in use on the system;
- the operating system in use, down to the level of specific version, the build number in use, and the history of security patches applied to it;
- IP addresses of Internet-connected computers;
- specific versions of systems administrator tools used;
- the security configuration of the operating system (e.g., whether certain services are turned on or off, or what antivirus programs are running); and
- the physical configuration of the hardware involved (e.g., what peripherals or computers are physically attached).

Note that none of these items of intelligence is easily available from satellite or aerial reconnaissance. As a general rule, a scarcity of intelligence regarding possible targets means that any offensive cyber operation launched against them can only be a “broad-spectrum” and a relatively indiscriminate or blunt attack. Such an attack might be analogous to the Allied strategic bombing attacks of World War II that targeted national infrastructure on the grounds that such infrastructure supported the war effort of the Axis. Substantial amounts of intelligence information about targets and paths to those targets are required if the operation is intended as a very precise one directed at a particular system. Conversely, a lack of

such information will result in large uncertainties about the direct and indirect effects of an operation and make it difficult to develop accurate estimates of likely collateral damage.

Active Defense

Defensive measures in cyber security seek to frustrate offensive operations taken against systems or networks. Passive defensive measures, such as hardening systems against penetration, facilitating recovery in the event of a successful offensive operation, making security more usable and ubiquitous, and educating users to behave properly in a threat environment, are important elements of a strong defensive posture.⁵ Nevertheless, for the defense to be successful, these measures must succeed every time an adversary attacks. The offensive operation need only succeed once, and an adversary who pays no penalty for a failed operation can continue with follow-on operations until it succeeds or chooses to stop. This places a heavy and asymmetric burden on a defensive posture that employs only passive defense.

If passive defense is insufficient to ensure security, what other approaches might help to strengthen one's defensive posture? One possibility is to eliminate or degrade an adversary's ability to successfully conduct offensive cyber operations. In that case, the operation is ultimately less successful than it might otherwise have been because the defender has been able to neutralize the operation in progress or perhaps even before it was launched.

A second possibility is to impose other costs on the adversary, and such a strategy is based on two premises. First, imposition of these costs reduces the adversary's willingness and/or ability to initiate or to continue an offensive operation. Second, knowledge that an operation will prove costly to one adversary deters others from attempting to conduct similar operations—and advance knowledge of such a possibility may deter the original adversary from conducting the offensive operation in the first place. There are many options for imposing costs on an adversary, including economic penalties such as sanctions, diplomatic penalties such as breaking of diplomatic relations, and even kinetic military actions such as cruise missile strikes. In-kind military action—a counteroffensive cyber operation—is also a possibility.

Both of these possible reactions—neutralization of an adversary's offensive operation and imposition of costs to the adversary for the operation—are often captured under the rubric of *active defense*. But note well—the

attempt to impose costs on an adversary that conducts offensive cyber operations might well be seen by that adversary as an offensive act itself. This may be especially true in the fog of cyber conflict, where who is actually doing what may be uncertain.

Evolving or Escalating Conflict

The phenomenon of escalation is a change in the level of conflict (where level is defined in terms of scope, intensity, or both) from a lower (perhaps nonexistent) to a higher level. Escalation is a fundamentally interactive concept in which actions by one party trigger other actions by another party to the conflict. Of particular concern is a chain reaction in which these actions feed off one another, thus raising the level of conflict to a level not initially contemplated by any party to the conflict. Escalation can occur through a number of mechanisms which may or may not be operative simultaneously in any instance.⁶ It includes four basic types: deliberate, inadvertent, accidental, and catalytic.

Deliberate escalation is carried out with specific purposes in mind. For example, a party may deliberately escalate a conflict from some initial level (which may be zero) to gain advantage, to preempt, to avoid defeat, to signal an adversary about its own intentions and motivations, or to penalize an adversary for some previous action. Offensive cyber operations—specifically, cyber attacks—are one of many possible military options for deliberate escalation.

Inadvertent escalation occurs when one party deliberately takes actions that it does not believe are escalatory but which are interpreted as escalatory by another party to the conflict. Such misinterpretation may occur because of incomplete information, lack of shared reference frames, or one party's thresholds or "lines in the sand" of which other parties are not aware. Communicating to an adversary the nature of any such thresholds regarding activity in cyberspace may be particularly problematic, even under normal peacetime circumstances.

For example, Nation A does X, expecting Nation B to do Y in response. But in fact, Nation B unexpectedly does Z, where Z is a much more escalatory action than Y. Or Nation A may do X, expecting it to be seen as a minor action intended only to show mild displeasure and that Nation B will do Y in response, where Y is also a relatively mild action. However, due to a variety of circumstances, Nation B sees X as a major escalatory action

and responds accordingly with Z, an action that is much more significant than Y. Nation A perceives Z as being way out of proportion and, in turn, escalates accordingly.

Accidental escalation occurs when some operational action has direct effects that are unintended by those who ordered them. A weapon may go astray to hit the wrong target; rules of engagement are sometimes unclear; a unit may take unauthorized actions; or a high-level command decision may not be received properly by all relevant units. It is especially relevant here that there is often greater uncertainty of outcome due to a lack of adequate intelligence on various targets when certain kinds of offensive cyber operations are employed.

Catalytic escalation occurs when some third party succeeds in provoking two parties to engage in conflict. For example, Party C takes action against Party A that is not traced to Party C and appears to come from Party B. Party A reacts against Party B, which then believes it is the target of an unprovoked action by Party A. The inherent anonymity of cyber operations may make “false-flag” operations easier to undertake in cyberspace than with kinetic operations.

Through such mechanisms, the escalatory dynamics of conflict show how a conflict, once started, might evolve. Of interest are issues such as what activities or events might set a cyber conflict into motion, what the responses to those activities or events might be, how each side might observe and understand those responses, whether responses would necessarily be “in-kind,” or how different kinds of states might respond differently.

Theories of escalation dynamics have been elaborated in the nuclear domain. But the deep and profound differences between the nuclear and cyber domains suggest that any theory of escalation dynamics in the latter would require far more than small perturbations in nuclear escalation dynamics theories, though such theories might be useful points of departure for developing new ones applicable to cyberspace. Some of these differences include the greater uncertainties in attribution of cyber actors, the broad proliferation of significant capabilities for cyber operations to a multitude of states and a variety of nonstate actors as well, and the inherent ambiguities of cyber operations compared to the very distinct threshold of nuclear weapons explosions.

To suggest some of the difficulties involved, consider the following scenarios:

- Nation Blue may believe it has been attacked deliberately by Nation Red, even though Red has not done so. Indeed, because of the ongoing

nature of various attack-like activities (e.g., hacking and other intrusions) against the computer systems and networks of most nations, Blue's conclusion that its computer systems are being attacked is certainly true. Attribution of such an attack is a different matter, and because hard evidence for attribution is difficult to obtain, Blue's government may make inferences about the likelihood of Red's involvement by giving more weight to a general understanding of Red's policy and posture toward it than might be warranted by the specific facts and circumstances of the situation. Evidence that appears to confirm Red's involvement will be easy to find, whether or not Red is actually involved. If Red is a technologically sophisticated nation (such as the United States), the lack of "fingerprints" specific to Red can easily be attributed to its technological superiority in conducting such attacks.

- An active defense of its systems and networks undertaken by Nation Red against Nation Blue could have significant political consequences. For example, even if Red had technical evidence that was incontrovertible (and it never is) pointing to Blue's government, Blue could still deny that it had launched such an attack—and in the court of world opinion, its denial could carry some credibility when weighed against Red's past assertions regarding similar issues. That is, Red's cyber attacks (counter-cyber attacks, to be precise) undertaken under the rubric of active defense may not be perceived as innocent acts of self-defense, even if they are. The result could be a flurry of charges and countercharges that would further muddy the waters and escalate the level of political tension and mistrust. The point at which a software agent for cyber attack is introduced or planted on an adversary's computer system or network is, in general, different from the point at which it is activated and begins to do damage. Blue (the nation being attacked) may well regard the hostile action as beginning at the moment Red's agent is planted, whereas Red may believe the hostile action begins only when the agent is activated.
- During periods of crisis or tension when military action may be more likely, it is entirely plausible that Blue would increase the intensity of security scans it conducts on its critical systems and networks. More intense security scans often reveal offensive software agents implanted long before the onset of crisis and that may have been overlooked in

ordinary scans, and yet discovery of these agents may well prompt fears that an attack is pending.

- The direct damage from a cyber attack is often invisible to outsiders. Without CNN images of smoking holes in the ground or troops on the move, an outside observer must weigh competing claims without tangible evidence one way or the other. Under such circumstances, the reputations of the different parties in the eyes of each other are likely to play a much larger political role.
- Nation Red plants software agents in some of Nation Blue's critical networks to collect intelligence information. These agents are designed to be reprogrammable in place—that is, Red can update its agents with new capabilities. During a time of crisis, Blue's authorities discover some of these agents and learn that they have been present for a while, that they are sending back very sensitive information to Red, and that their capabilities can be changed on a moment's notice. Even if no harmful action has yet been taken, it is entirely possible that Blue would see itself as the target of Red's cyber attack.

What follows are some speculations on some of the factors that might influence the evolution of a cyber conflict (see fig. 1).

Crisis Stability

Where kinetic weapons are involved, crisis stability refers to that condition in which neither side has incentives to attack first. Crisis stability is especially important for nuclear weapons, where the existence of an invulnerable submarine-based nuclear missile force controlled by Nation Blue means that Nation Red could not escape retaliation no matter how devastating a first strike it could launch. In terms of cyber weapons, there is no conceivable way for one nation to eliminate or even significantly degrade the cyber attack capability of another.⁷ But the question remains whether a second-strike cyber attack capability is the enabling condition for crisis stability in cyberspace.

A related question is that of incentives for preemption. Preemptive attacks by Red against Blue are undertaken to prevent (or at least blunt) an impending attack by Blue on Red. If Blue is planning a cyber attack on Red, a preemptive cyber attack on Blue cannot do much to destroy Blue's attack capability; at best, Red's preemptive attack on Blue might tie up Blue's personnel skilled in cyber operations. On the other hand, it is hard

Crisis Stability

- What is the analog of crisis stability in cyber conflict?
- What are the incentives for preemptive cyber attack?

Escalation Control and Management

- How can intentions be signaled to an adversary in conflict?
- How can cyber conflict between nations be limited to conflict in cyberspace?
- What thresholds of “line-crossing” activity might be created in cyberspace, and how might these be communicated to an adversary?
- How should cyber attack be scoped and targeted so that it does not lead an adversary to escalate a conflict into kinetic conflict?
- How can a modestly scoped cyber attack conducted by a government be differentiated from the background cyber attacks that are going on all of the time?
- How can the scale and scope of a commensurate response be ascertained?
- What confidence-building measures might actually reassure an adversary about a lack of hostile intent?

Complications Introduced by Patriotic Hackers

- How can “freelance” activities on the part of patriotic hackers be handled?

Incentives for Self-Restraint in Escalation

- What are the incentives for self-restraint in escalating cyber conflict?

Termination of Cyber Conflict

- What does it mean to terminate a cyber conflict?

Necessary Capabilities for Escalation Management

- How can national authorities exercise effective command and control of cyber forces in a rapidly evolving conflict environment?
- What is the scope and nature of national capabilities (e.g., technological, command and control, law enforcement/legal capabilities) needed to implement any approach to escalation management and conflict termination in cyberspace?
- How can each side obtain realistic assessments of one’s own or an adversary’s cyber state and condition (e.g., heavily or lightly damaged)?
- How might other resources/capabilities available to the United States be used to manage escalation of conflict in cyberspace?

Figure 1. Questions about escalatory dynamics of cyber conflict between nation-states

to imagine circumstances in which Red would realize that Blue were planning an attack, as preparations for launching a cyber attack are likely to be invisible for the most part.

A second relevant scenario is one in which Blue is planning a kinetic attack on Red. Intelligence information, such as photographs of troop movements, indicates preparations for such an attack. Under these circumstances, Red might well choose to launch a preemptive cyber attack with the intent of delaying and disrupting Blue's preparations for its own.

Signaling Intentions in Cyber Conflict

Nothing in the set of options above is specific to cyber conflict—such issues have been an important part of crisis management for a long time. But managing such issues may well be more difficult for cyber conflict than for other kinds of conflict. One reason is the constant background of cyber-attack activity. Reports arrive constantly of cyber attacks of one kind or another on US computer systems and networks, and the vast majority of these attacks do not have the significance of a serious cyber attack launched by a party determined to do harm to the United States. Indeed, the intent underlying a given cyber attack may not have a military or a strategic character at all. Organized crime may launch a cyber attack for profit-making purposes. A teenage hacking club may launch a cyber attack out of curiosity or for vandalism purposes.

Thus, if one nation wishes to send a signal to its cyber adversary, how is the latter to recognize that signal? Overtly taking credit for such an attack goes only so far, especially given uncertain communications in times of tension or war and the near certainty of less-than-responsible behavior on the part of one or both sides.

A dearth of historical experience with the use of serious offensive cyber operations further complicates efforts at understanding what an adversary might hope to gain by launching a cyber attack. In the absence of direct contact with those conducting such operations—sometimes even in the presence of such contact—determining intent is likely to be difficult and may rest heavily on inferences made on the basis of whatever attribution is possible. Thus, attempts to send signals to an adversary through limited and constrained military actions—problematic even in kinetic warfare—are likely to be even more problematic when cyber attacks are involved.

Determining the Impact and Magnitude of Cyber Response

If an adversary conducts a cyber attack against the United States, the first questions for US decision makers will relate to impact and magnitude. Such knowledge is necessary to inform an appropriate response. If, for example, the United States wishes to make a commensurate response, it needs to know what parameters of the incoming attack would characterize a commensurate response.

In many kinds of cyber attack, the magnitude of the impact of the first attack will be uncertain at first and may remain so for a considerable period of time. Decision makers may then be caught between two challenges—a policy need to respond quickly and the technical fact that it may be necessary to wait until more information about impact and damage can be obtained. These tensions are especially challenging in the context of active defense and active threat neutralization.

Decision makers often feel intense pressure to “do something” immediately after the onset of a crisis, and sometimes such pressure is warranted by the facts and circumstances of the situation. On the other hand, the lack of immediate information may prompt decision makers to take a worst-case view of the attack and, thus, to assume that the worst that might happen was indeed what actually happened. Such a situation has obvious potential for inappropriate and unintended escalation or kinetic response.

Transparency and Confidence-Building Measures

Where kinetic weapons are concerned, transparency and confidence-building measures such as adherence to mutually agreed “rules of the road” for naval ships at sea, prenotification of large troop movements, and noninterference with national technical means of verification have been used to promote stability and mutual understanding about a potential adversary’s intent.

Translating traditional transparency and confidence-building measures into cyberspace presents many problems. For example, generating forces in preparation for offensive cyber operations can be done essentially behind closed doors and with a small footprint, so evidence suggesting impending hostile action will never be evident, except with advance public notice. Thus, there is no reasonable analog for “notification of movement or massing of forces.” Because the success of offensive cyber operations is largely dependent on stealth and deception, reassurances of Nation Blue regarding the benign nature of any cyber activity observed, assuming it can be seen and attributed, ring hollow

to any parties that have a competitive or politically tense relationship with Blue. Traditional kinetic operations—those military operations on land, sea, and air—are easily distinguishable from most nonmilitary movements. By contrast, it is often difficult to distinguish between military and nonmilitary cyber operations, particularly between cyber attack and cyber exploitation. During a crisis, Blue may consider collecting intelligence on Red as stabilizing and thus lower the likelihood of mistaken escalation. Red may well interpret this as Blue preparing the battlefield as a prelude to attack.

These comments are not meant to suggest that all transparency or confidence-building measures for cyberspace are futile—only that applying traditional measures to cyberspace will be difficult, and new forms of conduct and behavior may be needed to promote transparency and build confidence.

Catalytic Cyber Conflict

Catalytic conflict as mentioned earlier refers to the phenomenon in which a third party instigates or seeks to escalate conflict between two other parties. These could be nation-states or subnational organizations such as terrorist groups. To increase confidence in the success of initiating a catalytic war, the instigator might attack both parties, seeking to fool each into thinking the other is responsible.

Because high-confidence attribution of cyber attacks under all circumstances is highly problematic, an instigator would find it relatively easy to deceive each party about the instigator's identity; thus, a double-sided catalytic attack may be plausible. Also, if a state of tension already exists between the two parties involved, leaders in each nation will be predisposed toward thinking the worst about the other, making them less likely to exercise due diligence in carefully attributing an attack. An instigator might consequently choose just such a time to conduct a catalytic cyber attack.

Complications Introduced by Patriotic Hackers

When traditional kinetic military operations are involved, it is generally presumed that the forces involved engage in armed conflict only at the direction of the cognizant government, only by its authorized military agents, and specifically, not by private groups or individuals. That is, governments maintain their armed forces to participate in armed conflict under the government's direction.

But in the Internet era, it is necessary to consider that nonstate actors may become involved in conflict. During times of conflict (or even tension) with another nation, some citizens may be motivated to support their country's war effort or political stance by taking direct action in cyberspace (see fig. 2). Such individuals—often known as hacktivists or patriotic hackers—are private citizens with some skills in the use of cyber attack weapons, and they may well launch cyber attacks on the adversary nation on their own initiative; that is, without the blessing and not under the direction or control of the government of that nation.

A number of incidents of privately undertaken cyber attacks have been publicized:

- Immediately after the start of the second intifada in Israel in late September 2000, Palestinian and Israeli hackers conducted a variety of cyber attacks on each other's national web presences on the Internet.⁸
- Following the 2001 incident between the United States and China in which a US EP-3 reconnaissance aircraft collided with a Chinese F-8 interceptor, both Chinese and American hackers attacked the web presence of the other nation. In both cases, attacks were mostly aimed at website defacement and denial of service.⁹
- In the wake of the May 1999 bombing by the United States of the Chinese embassy in Belgrade, the US National Infrastructure Protection Center issued an advisory (NIPC Advisory 99-007) noting "multiple reports of recent hacking and cyber activity directed at U.S. government computer networks, in response to the accidental bombing of the Chinese embassy in Belgrade. . . . Reported activity include[d] replacing official web pages with protest material and offensive language, posting similar language in chat rooms and news groups, and denial of service email attacks."¹⁰
- American hackers have been known to attack jihadist websites. For example, an American was reported by *Wired* magazine to have hijacked www.alneda.com, a widely used website for jihadist recruitment.¹¹ His motive for doing so was said to be a decision made after the September 11 attacks: "I was going to use every skill I had to screw up the terrorists' communication in any way I could."
- Russian hackers are generally reported to have been responsible for the cyber attacks on Estonia in 2007 and Georgia in 2008.¹²

Allen and Demchek generalize from experiences such as these to predict that future conflicts between nations may involve:

- Spontaneous attack action in cyberspace by "patriots" on each side.
- Rapid escalation of these actions to a broad range of targets on the other side—because hacktivists are interested in making a statement, they will simply attack sites until they find vulnerable ones.
- Involvement of sympathetic individuals from other nations supporting the primary antagonists.

Figure 2. Hacktivism during international conflict and tension. Adapted largely from Patrick D. Allen and Chris C. Demchak, "The Palestinian-Israeli: cyberwar" [*sic*], *Military Review*, March–April 2003.

The actions of these patriotic hackers may greatly complicate escalation management. Such actions may be seen by an adversary as being performed under the direction, blessing, tacit concurrence, or tolerance of the state and therefore are likely to be factored into the adversary's assessment of the state's motives and intent. The state's efforts to suppress patriotic hackers may be seen as insincere and are likely to be at least partially unsuccessful as well. In a worst-case scenario, actions of patriotic hackers during times of tension may be seen as an officially sanctioned cyber first strike, even if they have not acted with government approval or under government direction.

Yet another complication involving patriotic hackers is the possibility that they might be directed by, inspired by, or tolerated by their government but in ways in which the government's hand is not easily visible. Under such circumstances, hostile acts with damaging consequences could continue to occur with corresponding benefits to the nation responsible despite official denials. At the very least, the possibility that patriotic hackers may be operating could act as a plausible cover for government-sponsored cyber attacks, even if there were in fact no patriotic hackers doing anything.

Incentives for Self-Restraint in Escalation

One set of incentives is based on concerns about an adversary's response to escalation. Understanding this set of incentives is necessarily based on a sense of what kinds of offensive cyber actions—whether cyber attack or cyber exploitation—might be mistaken for cyber attack and might lead to what kinds of adversary responses, either in cyberspace or in physical space. In this regard, an essential difference between cyber attack and the use of a nuclear, chemical, biological, or space weapon is readily apparent—the initial use of any nuclear, chemical, biological, or space weapon, regardless of how it is used, would constitute an escalation of a conflict under almost any circumstances. By contrast, whether a given cyber attack, or conventional kinetic attack for that matter, would be regarded as an escalation depends on the nature of the operation—the nature of the target(s), their geographical locations, or their strategic significance.

A second set of incentives is based on concerns about blowback—the possibility that a cyber attack launched by the United States against Nation B's computers might somehow affect US computers at a later time. Understanding the likelihood of blowback will require a complex mix of technical insight and intelligence information.

Deescalation and Conflict Termination

Conflict termination presumes the existence of an ongoing conflict to which the participants desire an end. It requires several elements, including:

- a reliable and trustworthy mechanism that can be used by the involved parties to negotiate the terms of an agreement to terminate a conflict,
- a clear understanding on all sides about what the terms of any agreement require each side to do,
- assurance that all parties to an agreement will adhere to the terms of any such agreement, and
- capabilities for each party that can insure all entities taking action on behalf of that party adhere to the terms of any such agreement.

In the cyber environment, these elements may be problematic. National leaders and their representatives will almost certainly be communicating with each other through electronic channels, the reliability of which may be questionable in certain kinds of cyber conflict. A cease-fire agreement in cyberspace presumes each side can know that the other has stopped hostile activity in cyberspace. However, ambiguity and technical limitations create problems. Nation Blue may conduct cyber exploitations seeking to verify that Nation Red is standing down in cyberspace. Red may interpret these operations as prelude to Blue's continuing an attack campaign against it. Patriotic hackers of Blue may press onward against Red even though both Red and Blue have themselves agreed to a cyber cease-fire. During conflict, there is no reason to assume the cessation of continuing cyber operations conducted by others who are not part of the conflict (e.g., criminals). In some cases, ongoing offensive operations by these third parties may be mistakenly attributed to Red or Blue. The two nations may differ in their interpretation of key concepts. What activities constitute an "attack" in cyberspace, or what evidence should be used to determine if an attack is occurring? Differing interpretations and inadequate technical capabilities may impede understanding. For kinetic military forces, a variety of technical means (e.g., photoreconnaissance aircraft and satellites, ocean-scale sonar arrays) are capable of monitoring movements of military personnel and equipment. Most importantly,

these means operate from outside territory controlled by an adversary and provide information that is generally regarded as reliable. But because the footprint of cyber forces is so small, movement of adversary forces can take place without signatures that can be externally observed. Based on precedents in kinetic conflict, it is plausible that nations seeking a cease-fire in a cyber conflict would ask for the deactivation of these hostile agents. To comply with such a request (not an unreasonable one in the context of a cease-fire), these nations will need to maintain cyber “demin-ing” capabilities regarding the offensive software and/or hardware agents they implant into adversary systems, networks, and infrastructure. For example, they will need to keep track of where these agents are implanted or be able to communicate with them to disarm them—a capability that may rule out offensive agents that operate in a fully autonomous manner.

Each party will naturally have concerns about its adversary’s commitment to adhere to the terms of a cyber cease-fire, especially in the aftermath of a conflict. On what basis would Blue’s government believe a claim by Red that it was indeed complying with the terms of a cease-fire? How much would Red tell Blue about system and network penetrations it had made, knowing such information might be used to prosecute an attack or defend more effectively against Red? The availability of effective ways to address the issues described above is almost certainly one aspect of being able to manage conflict termination in cyberspace.

Analysts sometimes raise the issue of how the United States might deter escalation when it has more at stake in cyberspace than its adversaries. The first point to consider is that deterrence of cyber attack does not necessarily entail a threat to respond through cyberspace against an adversary’s cyber assets, and when non-cyber threats against an adversary’s non-cyber assets are considered, the calculus of deterrence may well be different. For example, kinetic weapons can, in principle, be employed against valuable physical military targets. Although the threshold for such a response may well be higher, an adversary would still have to consider the possibility of a non-cyber response to any attack. Consistent with this point, US policymakers have always noted that the United States reserves the right to respond appropriately in a time, place, and manner of its own choosing. In addition, concerns over blowback may deter an adversary. If an adversary’s interests are entangled with those of the United States, it may be deterred from taking actions that might harm US interests because of concerns that one ultimate effect of such actions would be to harm the

adversary's interests. For example, a nation that is owed a great deal of money by the United States might well be unlikely to conduct an attack that undermines its financial stability.

Lastly, many analysts note that deterrence is a psychological phenomenon and that threats of retaliation must be focused on assets that an adversary holds dear and values highly. In principle, what an adversary—or more precisely, an adversary decision maker—holds dear can span a wide range, from personal to national (e.g., tools of national power). In the category of personal assets are financial entities (e.g., a leader's bank accounts could be drained), reputation (e.g., a scandal in a policymaker's past might be revealed), and close friends and relatives (e.g., the interests of such individuals could be compromised). Such assets are not typically considered in a traditional military context—but nontraditional approaches to deterrence may well be needed to deal with the nontraditional threats that cyber attacks pose.

The approaches described above may be most useful in deterring hostile cyber operations intended to achieve large-scale effects. They are unlikely to be useful in deterring operations intended to achieve smaller effects, because smaller effects by definition do not cause maximum pain for either side. Put differently, the argument that the United States has more at risk in cyberspace than its adversaries is simply not relevant when the amount of damage that can be done (by definition) is small.

Kinetic Escalation

Issues of escalation and conflict termination in cyberspace are complicated by the fact there may be cross-domain linkages. Although conflict might, in principle, be limited to hostile operations in cyberspace alone, there is no reason this is necessarily so, and policymakers must contemplate the possibility that conflict in cyberspace might spill over into physical space, and might even lead to kinetic actions.

For example, if national command authorities decide to retaliate in response to a cyber attack, an important question is whether retaliation must be based on a “tit-for-tat” response. Assuming the perpetrator of a cyber attack is known to be a hostile nation, there is no reason in principle the retaliation could not be a kinetic attack against the interests of that hostile nation. Allowing a kinetic response to a cyber attack expands the range of options available to the victim. An extreme case is, in the event of a cyber attack of sufficient scale and duration that it threatens the nation's

ability to function as a modern society, the attacked nation might choose to respond with kinetic force. On the other hand, the use of kinetic operations during an ostensibly cyber-only conflict is an important threshold. Nations involved in a cyber-only conflict may have an interest in refraining from a kinetic response—for example, they may believe kinetic operations would be too provocative and might result in an undesired escalation of the conflict.

In addition, the logic of offensive cyber operations suggests that such operations are likely to be most successful when the initiator of these operations has the time to gather intelligence on likely targets—such intelligence gathering is obviously time-limited once overt kinetic conflict breaks out.

If understanding the dynamics of cyber-only conflict is difficult, understanding the dynamics of cyber conflict when kinetic operations may be involved is doubly so. To the extent national decision makers have incentives to refrain from conducting offensive operations that might induce a strong kinetic reaction, the obvious approach would be to conduct cyber attacks that are in some sense smaller, modest in result, targeted selectively against less-provocative targets, and perhaps more reversible. The similarity of such an approach to escalation control in other kinds of conflict is not accidental, and it has all of the corresponding complexities and uncertainties.

In keeping a cyber conflict from escalating into physical space, it is important to think about “lines in the sand” beyond which one side warns another not to cross. For example, it is reported that during the first Gulf War, the United States regarded Iraqi use of chemical weapons against US forces as one such threshold of unacceptable activity, one that might well provoke the use of US nuclear weapons against Iraq. When only traditional kinetic forces are involved, lines in the sand might be the use of certain weapons, attacks on or damage to certain targets, movement or placement of armed forces beyond certain geographical lines, and so on. Cyber analogs to these thresholds are hard to construct. Describing a class of cyber weapon whose mere use would be wholly unacceptable is hard to imagine, since there are no real cyber analogs to true weapons of mass destruction where even a single use of a WMD qualitatively changes the landscape of kinetic conflict. And in cyberspace, what is the analog of a geographical border beyond which cyber weapons may not be placed?

Perhaps the most promising analog is the notion of specific targets that might be placed off limits—cyber attacks on such targets could, in principle,

be deemed unacceptable. One class of off-limits targets might be cyber assets associated with truly critical infrastructure, such as the bulk power grid or the banking and financial system. But as any bank executive will confirm, some of these targets are under attack quite frequently—so attacks that do not cause large amounts of damage or loss probably should *not* qualify as crossing the threshold of unacceptability. There is also the question of being able to assign *political* responsibility to some perpetrator for the conduct of a successful large-scale attack on some off-limits target—a question whose answer may be in doubt, given the difficulties of rapid attribution of a cyber attack. Finally, one might well ask how a cyber asset would be positively identified as being associated with the bulk power grid or the banking and financial system. Would we provide a computer-readable identification tag on every such computer? Such a tag might make these targets obvious to other parties wishing to do us harm.

Even presuming that the United States could identify specific thresholds, such information would need to be communicated clearly to an adversary. Such communication is difficult even in scenarios of traditional military conflict, and all of these difficulties obtain in the cyber context. But it is worth observing that because cyber conflict is fundamentally based on deception, persuading an adversary to believe any US statement about what is off-limits may be particularly challenging.

The Political Side of Escalation

Despite the focus of the discussion above on escalation dynamics from a primarily military standpoint, escalation dynamics inevitably have a political and psychological component that must not be overlooked. For example, the discussion of active defense above pointed out that US cyber attacks undertaken under the rubric of active defense may not be perceived by others as innocent acts of self-defense, even if they are intended as such. While both sides in most conflicts claim they are acting in self-defense, cyber conflicts are a particularly messy domain in which to air and judge such claims.

Another possible misperception may arise from intelligence-collection activities that might involve cyber-attack techniques. The discussion above noted the problems of misperceiving exploitation as a prelude to continuing cyber operations during a cease-fire. But the problem is broader than that—during conflict or in the tense times that often precede con-

flict, the needs for current intelligence on the adversary are particularly acute. Knowing what the adversary is doing and the scope and nature of its future intentions are very important to decision makers, and the need to collect such intelligence will almost certainly result in greater pressures to use the entire array of available intelligence-gathering techniques—including techniques of cyber exploitation. If the adversary is unable to distinguish between an offensive operation for exploitation and one for attack—an outcome that seems all too likely—a cyber exploitation may run the risk of being perceived as part of an imminent attack, even if this is not the intent of decision makers.

Finally, it seems likely that escalation issues would play out differently if the other nation(s) involved are or are not near-peer competitors. Escalation to physical conflict is less of a concern to the United States if the nation has weak conventional forces and/or is a nonnuclear state. But a nation with nuclear weapons, or even strong conventional forces in a position to inflict significant damage on US allies, is another matter entirely. Relationships with such states may well need to be explicitly managed, paying special attention to how escalation may be viewed, managed, and controlled, and most importantly, how miscalculation, misperception, or outright error may affect an adversary's response.

Dynamics such as these suggest that factors other than the ones dictated by military or legal necessity play important roles in escalation dynamics, if only because they can strongly affect the perceptions of decision makers on either side.

The Future of Escalation Dynamics

The issues of escalation dynamics, conflict termination, and cross-domain linkages in cyberspace play out against a rapidly changing technological, policy, and geopolitical environment. The substrate of cyberspace—computing and communications technology—is characterized by change on a timescale much shorter than the planning horizon for traditional military acquisitions and planning. Upgrades notwithstanding, major weapons platforms are expected to serve for decades, while the information technology environment changes rapidly in a few years. The growing use of cloud computing is a further—and potentially disruptive—change in possible computing platforms and may require new concepts for assigning responsibility for cyber operations. Mobile computing may present

opportunities for determining device location as well as being the enabling technology for many new users of cyberspace. IT will be increasingly embedded, ubiquitous, and connected within all elements of modern society, potentially increasing vulnerabilities to all manner of societal functions. The result is that operational concepts for escalation management must take into account a rapidly evolving set of targets and offensive and defensive capabilities.

In most traditional domains of conflict, US military doctrine has been based on the establishing dominance—that state in which friendly forces have maximum freedom of action and adversary forces have minimal freedom of action. But in the cyber domain, this presumption is not sustainable—and senior US military leaders are beginning to speak publicly about this point.¹³ Much of the traditional US approach to escalation control is based on the ability of friendly forces to establish dominance at any level of conflict on the premise that an adversary would not choose to escalate if, at the higher level of conflict, it could not hope to prevail.

Nation-states are increasingly concerned about the risks inherent in involvement in cyberspace. Even apart from the protection of critical national infrastructure and military assets, various nations express deepening worries about traditional criminal activity in cyberspace, protection of intellectual property, and increased connectedness for political movements that may pose a threat to government interests and stability.

Nonstate actors are increasingly important players in cyberspace. Multinational corporations and organized crime syndicates, for example, all have some nontrivial capability to conduct offensive operations in cyberspace to further their interests, and even small groups of individuals can have a large impact by exploiting certain characteristics of cyberspace (e.g., WikiLeaks).

Although existing theories of escalation dynamics and conflict termination may serve as useful points of departure, what is understood very poorly today is how these theories may apply in cyberspace. In the future, finding ways to manage cyber conflict will be even more intellectually challenging than it was for traditional conflict. **SSQ**

Notes

1. The lag time between dissemination of a security fix to the public and its installation on a specific computer system may be considerable, and it is not always due to unawareness on the part of the system administrator. It sometimes happens that the installation of a fix will cause an application running on the system to cease working, and administrators may have to weigh

the potential benefit of installing a security fix against the potential cost of rendering a critical application nonfunctional. Adversaries take advantage of this lag time to exploit vulnerabilities.

2. A zero-day attack is a previously unseen attack on a previously unknown vulnerability. The term refers to the fact that the vulnerability has been known to the defender for zero days. (The adversary has usually known of the attack for a much longer time.) The most dangerous is a zero-day attack on a remotely accessible service that runs by default on all versions of a widely used operating system distribution. This type of remotely accessible zero-day attack on services appears to be occurring less frequently. In response, a shift in focus to the client side has occurred, resulting in many recent zero-day attacks on client-side applications. For data and analysis of zero-day attack trends, see Daniel Geer, “Measuring Security,” *Dan@Geer.org*, 278–87, <http://geer.tinho.net/measuringsecurity.tutorialv2.pdf>.

3. An adversary computer or network may not necessarily be owned and operated by the adversary—it may simply support or be used by the adversary.

4. For purposes of this article, the term *attribution* is used to refer to the identification of the party to which political responsibility should be assigned for the cyber operations that harm the interests of the target. This qualifier is necessary because the entity “responsible” can also be the machine(s) involved in the operation or the specific human beings who took specific actions (at a keyboard) to launch the operation. One of these other meanings may be more relevant, depending on the purposes for which attribution is sought. For more discussion of this point, see David D. Clark and Susan Landau, “Untangling Attribution,” *National Security Journal*, 16 March 2011, <http://harvardnsj.org/2011/03/untangling-attribution-2/>, as well as William Owens, Kenneth Dam, and Herbert Lin, *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities* (Washington: National Academies Press, 2009), chap. 2.

5. The broad topic of how to improve passive cyber defenses and enhance resilience of US computer systems and networks is addressed in a variety of National Research Council (NRC) reports on this topic: *Computers at Risk*, 1991; *Information Technology for Counterterrorism*, 2003; *Cybersecurity Today and Tomorrow: Pay Now or Pay Later*, 2002; *Realizing the Potential of C4I: Fundamental Challenges*, 1998; *Trust in Cyberspace*, 1999; and *Toward a Safer and More Secure Cyberspace*, 2007, all authored by the NRC and published by National Academies Press, Washington, DC. Other important reports include President’s Information Technology Advisory Committee, *Cyber Security: A Crisis of Prioritization* (Washington: National Coordination Office for Information Technology Research and Development, February 2005); and Commission on Cyber Security for the 44th Presidency, *Securing Cyberspace for the 44th Presidency* (Washington: Center for Strategic and International Studies, 2008).

6. This taxonomy is based mostly on *Dangerous Thresholds: Managing Escalation in the 21st Century* (Santa Monica, CA: RAND, 2008), though the RAND discussion is silent on escalation in cyberspace per se.

7. Even in the case of a nuclear EMP attack directed against electronic equipment in another nation, there is no reason to assume that all of that nation’s cyber-attack capabilities are necessarily resident within its boundaries. Because cyber attacks can originate from anywhere, some cyber attack capabilities may have been deployed in other nations—indeed, some attack agents may already have been clandestinely deployed in US systems.

8. “Cyberwar Also Rages in Mideast,” *Associated Press*, 26 October 2000, <http://www.wired.com/politics/law/news/2000/10/39766>.

9. Michelle Delio, “A Chinese Call to Hack U.S.,” *Wired*, 11 April 2001, <http://www.wired.com/news/politics/0,1283,42982,00.html>.

10. Available at <http://www.merit.edu/mail.archives/netsec/1999-05/msg00013.html>.

11. Patrick Di Justo, "How Al-Qaida Site Was Hijacked," *Wired*, 10 August 200, <http://www.wired.com/culture/lifestyle/news/2002/08/54455>.
12. "Expert: Cyber-Attacks on Georgia Websites Tied to Mob, Russian Government," *Los Angeles Times*, 13 August 2008, <http://latimesblogs.latimes.com/technology/2008/08/experts-debate.html>.
13. For example, RADM William Leigher, deputy commander of the US Navy Cyber Command, was recently quoted as saying that "Unlike the physical domain, achieving dominance [in the cyber domain] may be impossible." Amber Corrin, "Dominance in Cyberspace Might not be Possible," *Defense Systems*, 27 January 2011, <http://defensesystems.com/articles/2011/01/27/afcea-west-cyber-warfare-panel.aspx>.

Sharing the Cyber Journey

Suzanne M. Vautrinot, Major General, USAF

0620 ZULU (1120 PDT): Based on remotely piloted aircraft (RPA) surveillance, special operations forces prepare to enter a village that contains a high-value target (HVT).

0630 ZULU: The mission commander in the joint operations center monitors the HVT and surrounding village activity via real-time video feed from the Predator aircraft.

0632 ZULU: The mission commander loses visual surveillance of the current operation.

- Did a civilian system administrator in California pull a circuit offline to perform routine maintenance?
- Did a highway construction crew in Florida cut a fiber-optic cable during excavation?
- Did an adversary nation inject malicious software, preventing operation of the common operating system display?
- Did lightning take out a transformer in Nevada and cut off power to the data transmission system?

0635 ZULU: Forces reach preposition points and stand by for mission authorization.

06?? ZULU: The mission commander aborts the mission due to lack of situational awareness.

As the forces hunker down, the entire command and a global support structure hit afterburner in an attempt to determine (1) what happened to cause the loss of visual contact, (2) can it be recovered, and (3) will it be in time to achieve the intended mission?—a situation with seemingly infinite causality, demanding action in finite moments.

Maj Gen Suzanne M. “Zan” Vautrinot is commander, Twenty-fourth Air Force, and commander, Air Force Network Operations, Lackland AFB, Texas. She is also the commander, AFCYBER, and is responsible for providing combatant commanders with trained and ready cyber forces to plan and conduct cyberspace operations. General Vautrinot is a 1982 US Air Force Academy graduate who has served in various cyber operations, plans and policy, strategic security, space operations, and staff assignments.

The author gratefully acknowledges the contributions of Capt Jeffrey A. Martinez and Capt Matthew R. Kayser in preparing this article.

While the operations center staff check their equipment, computer maintainers in dozens of locations check for indicators of hardware or software system failure; civil engineers evaluate power, chillers, and HVAC systems operation; network operators across the globe search for dropped fiber connections; satellite operators work to verify communication and data feeds; spectrum analysts look for jamming indications; intelligence analysts dive into indications of potential adversary action; weather experts evaluate scintillation—all while mission commanders check their watches.

One operation, one mission, yet it requires a myriad of extraordinary experts—each unique and each integral to an RPA operation that depends on well over a hundred individual commercial and military network connections, dozens of integrated hardware systems, miles of fiber-optic cable, significant satellite bandwidth, and millions of lines of software code. Welcome to the cyber domain: an environment of intellect, integration, and, for good as well as ill, complex interdependency.

The scenario described above could affect equally any military weapon system or mission. In the vast majority of cases, these network dependencies are not well documented, the real-time status of network systems is not automated or transmitted, the supporting infrastructure is diverse and aging, the investigation remains essentially manual, and the fingers generally point to the “distant end,” located in the vicinity of Valhalla. One might conclude poor performance, inadequate resourcing, or perhaps poor design, but the dynamics simply reflect the way cyber has rapidly emerged—in our equipment and in our collective psyche.

Historically, technology was leveraged to improve performance of each weapon system relative to the environment in which it must operate. That environment was governed by Mother Nature, and our ability to fly through, dive beneath, breathe without, orbit above, or move undetected was achieved by creating systems that overcame environmental limitations. Each new technology was ingeniously integrated into our ground, sea, air, and space systems to gain capability. By leveraging communications, computers, networks, and information technology, we improved the capabilities of each existing system while also making them dependent on a new environment—a man-made cyber environment. The acute dependency was unintentional, and like our legacy networks, it grew with the best of intentions and a dearth of strategic design.

A strategic discussion on cyber has become more than a DoD activity; it is now a national imperative. As Malcolm Gladwell might say, we are at

a tipping point. Relative to cyber technologies, do we continue to bolt on or should we bake in? Regarding cyberspace as a man-made environment, do we simply respond to changes or work with our civil sector counterparts to alter the environment to our collective advantage? As we leverage the technologies associated with cyberspace, we have an opportunity to constantly create and re-create our environment—to design the future.

Leveraging the Past, Innovating the Future

Every generation stands on the many shoulders of greatness that preceded it. For military leaders and as part of our Air Force heritage, flying faster, turning tighter, launching further, viewing in more detail, and arriving with greater precision all align with a tradition of innovating beyond the heritage left by revered forefathers. The world we face today is significantly different from that of our predecessors. From a military perspective, the most formidable changes do not just involve enhancing the physical attributes of our weapon systems or incrementally adjusting the traditional methods of employing those weapon systems. The distinction is that now we can leverage the virtual, and the implications are boundless.

We did not arrive at this point overnight. For decades, leaders in engineering, cryptology, computer science, information technology, and many other contributing disciplines expanded and then integrated these technologies. Yet, although the technical disciplines were varied, the application of cyber now follows a path similar to air, sea, and space in their early stages. Akin to the Wright Flyer's relationship to the F-35, mainframes, and eventually personal computers, were the harbingers of our cyber capabilities. Continued platform development led to aircraft being used as a ground force and intelligence enabler during Army Air Corps operations. Similarly, integrated networks enabled the rapid dissemination of information for defense and intelligence operations. Code-breaking and cryptology applied to secure communications foreshadowed today's cyber information assurance and exploitation capabilities.

Airpower eventually emerged as both a supporting element and a formidable alternative to traditional land and sea forces. The application of cyber capability to enable ground, sea, air, and space operations continues to accelerate, but as with airpower, we should similarly expect cyber to emerge as a strategic alternative.

To advance cyber toward this strategic alternative, Twenty-fourth Air Force (24 AF) was established as a war-fighting numbered air force focused on full-spectrum cyberspace operations. It operates under three distinct roles: Air Forces Cyber (AFCYBER), the USAF cyber component force provider to combatant commanders (COCOM) through US Cyber Command; AF Network Operations (AFNetOps), the operator and defender of the Air Force portion of the DoD network; and 24 AF, the organize, train, and equip lead for USAF cyber personnel. Since both the AFNetOps and 24 AF functions oversee USAF-specific mission areas, they report to Air Force Space Command (AFSPC); in the AFCYBER role, they report directly to US Cyber Command and provide capabilities at the operational level to the joint war fighter.

Currently, we have a reactive defense posture that is outdated and manpower intensive. Our heterogeneous architecture, composed of legacy infrastructures, is difficult to maintain and provides limited situational awareness across the networks. With a steady topline cyber funding amount, as depicted in figure 1, every dollar spent toward protecting our networks needs to move us toward a more homogeneous and centralized

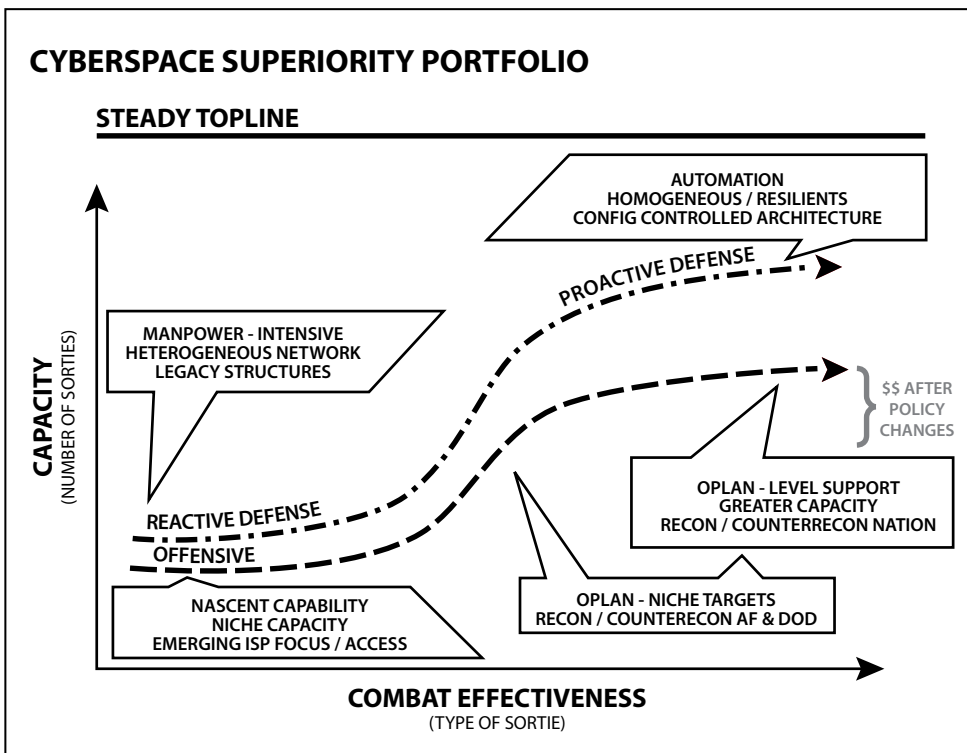


Figure 1. Cyberspace investment challenge

architecture that can reap the benefits of automation. Future investments must reflect advancement toward automation and resilient architectures so the efficiencies gained in manpower can increase the capacity of a skilled technical workforce.

We are at a nexus regarding future cyberspace operations providing for the national defense. For the Air Force to fulfill its commitment to providing global reach, global power, and global vigilance, it must do what Airmen have always done—innovate. To accomplish these goals, we have developed three integrated strategies: deliver a robust, defensible, trusted network; operationally leverage cyberspace capabilities; and build and deliver combat power. The remainder of this article is organized around an Air Force leadership dialog and Airmen's fulfillment of these strategies.

Deliver a Robust, Defensible, Trusted Network

The RPA exemplar applies equally to every military service member's ground, sea, air, or space operations; to their civilian counterparts' corporate business; and to local, state, or federal government activities. Each requires assurance that the networks, the multifaceted environment on which they are now so dependent, can be trusted to enable mission success.

Cyberspace is not simply the Internet; rather, it is a network of interdependent information technologies including the Internet, telecommunications networks, computer systems, and embedded processors. Its use has become ubiquitous within every public, industrial, academic, and military organization. Individually and collectively, we have increased productivity, interaction, performance, and efficiency by use of and by reliance on cyberspace. We "face-time" with friends and family, we pay bills via bank websites, parents monitor home security while away, and troops use social media to stay connected to home. Most importantly for this conversation, the nation and the Air Force have increased weapon system performance, extended operational capabilities, and enhanced command and control by leveraging cyberspace. Yet, as with all things yin, there is a yang. The dark side leverages this common ground to steal, compromise, degrade, or destroy information; disrupt networks or communications; or deny service. In military terms, cyberspace is a contested environment. Hactivists, cyber criminals, terrorists, and adversarial nations are active in cyberspace networks across the globe; our military networks are no exception. DoD networks are probed millions of times per day. In a typical

week, the Air Force blocks roughly two billion potential threats and denies two million phishing or spam e-mails. Armed with an understanding of the growing threat to and our dependency on the network, Air Force leaders directed a service-wide migration to a more defensible network—creating the AFNet migration and facilitating a “defense-in-depth” alignment. Helping create this defensible construct, AFSPC, through its subordinate units at 24 AF and the Air Force Network Integration Center, is reorganizing and reequipping to address the limitations resident in current Air Force heterogeneous network architecture and the underlying technologies. What is meant by “heterogeneous” network? We have many variances in hardware, configuration, and software licensing. As the network expands, updating and maintaining various systems becomes problematic. Inevitably, devices are not properly or consistently configured, and vulnerabilities arise. Moreover, the ability to discern the “root cause” of network issues requires significant time and resources to first understand the configuration, then find and address the underlying problems.

The process of moving from this dispersed, installation-managed network architecture to a single, homogeneous, and centrally managed Air Force network, called the AFNet, is the number one cyberspace initiative in the Air Force. Originally, the AFNet migration consisted only of consolidation of individual base active directory “trees” into a single Air Force active directory tree. Now the term has evolved into a broader concept involving all the necessary steps to move to a single Air Force network. Industry counterparts like AT&T preceded us in this endeavor, applying significant up-front capital and draconian change management. Their conclusion, and ours, is that without the initial homogeneity, we cannot implement the necessary sensors and automation to strengthen and defend network operations at the scale required for a global industry or military operations.

The first step was to realign AF network interfaces through a small number of gateways, thereby increasing visibility of network traffic as it moved into and among various organizations. This allows Air Force operators to observe patterns of (network) behavior and respond to anomalous activity. That response can include notification of other service and DoD-level operations centers (notably the joint operations center for US Cyber Command), implementing passive defenses within the AFNet, conducting forensics, reverse-engineering software, and supporting law enforcement and/or intelligence professionals in tracing the sources and potential implication of

intrusions. The vast majority of this work remains appropriately invisible to network users; nevertheless, it is foundational to a defensible network.

The second step of migration involves consolidation of each individual base's active directory structure into a single Air Force active directory tree. Simply put, active directory enables a centralized approach for network management and security. It provides services that authenticate and authorize users, assigns and enforces cyber policies, and simplifies updating computers. This will enable a simpler, more automated approach to managing the Air Force's e-mail and SharePoint applications. In addition, it will allow shutdown of the legacy systems at each base. Airmen at all levels and every base continue to rise to the challenge, and to date, roughly a quarter of all locations have migrated, with a targeted completion in FY-13. Migrating the entire Air Force population of roughly 850,000 personnel at over 400 locations will result in a much more defensible construct that aligns the Air Force leadership vision with the guidance and intent of US Cyber Command: to provide a more secure and, ultimately, operational platform.

There are many advantages to this AFNet migration, the most important being the opportunity to now increase sensing, automation, and situational awareness. In the Central Command Combined Air Operations Center, walls are filled with screens depicting operational status and battle-field video feeds for real-time analysis and decision making. The corresponding cyber information to depict network operational status and enable real-time analysis does not currently exist, nor was it possible prior to the rearchitecting of the AFNet. Operators in the 624th Operations Center, 24 AF's command and control unit, manually perform the task of data synthesis after distant-end units enter status information into the system. There is no common operating picture of activity across our networks, making it more difficult to assess and respond to the threat environment. Yet, there are innovators: cyber professionals from many career fields who daily apply capabilities and leverage new tactics, techniques, and procedures to successfully provide mission assurance, threat detection and response, and network operations and defense. The capabilities for sensing the status and automating operational activities will continue to expand, and so must the capacity elements necessary to reach and execute full-spectrum cyber operations globally. Migration to a single architecture provides the opportunity for Air Force-wide network situational

awareness—an awareness that enables robust, defensible, and trusted air, space, and cyber operations.

When designers of major weapon systems build cyber technologies into their programs, they fail to integrate them with the Air Force network. Frequently, these systems introduce cyber vulnerabilities into the network that cannot be patched or updated using established capabilities and processes. Networks cannot just be the domain of cyber folks; they must be central in the development and operation of every weapon system for design and connection interfaces. This requires application and enforcement of network standards for any weapon system that uses the Air Force network.

In that pursuit we are striving to increase awareness of rapid technological advances and best practices through partnerships with academia, industry, sister services, and government agencies. General Alexander outlined in his recent remarks to the Senate Armed Services Committee that, in his view, there are three key players that make up a cross-government team to mature and implement an effective cyber strategy for the nation: the Department of Homeland Security, the Federal Bureau of Investigation, and the DoD/intelligence community/National Security Agency/USCYBERCOM. Through USCYBERCOM, we have teamed with cyberspace law enforcement counterparts: leaders like Steve Shirley at the DoD Cyber Crime Center, and the OSI to share information on current threats and tactics as well as leverage their unique forensics expertise. Via 24 AF and the Air Force Computer Emergency Response Team (CERT), the USAF participates in the Defense Industrial Base Initiative, an agreement with over 30 industry partners, including many of the larger corporations in this country, to collaborate with the Departments of Defense and Homeland Security to share sensitive threat information and thereby improve the collective cyberspace defense. Moving forward, we will continue to leverage the great capacity and unique capabilities of not only 24 AF and Air Force Space Command but also the expertise of Airmen in our intelligence, law enforcement, and engineering development communities.

The Air Force utilized partnerships with Department of Energy and university national laboratories, like Lawrence Livermore National Laboratory, to deliver a network defense system in the early 1990s. We continue to develop and expand those core relationships today. We are working with Lawrence Livermore to field a network situational awareness capability that is being used by other government organizations. These channels for coop-

eration increase the flow of information and create a higher level of awareness across all levels of academia, industry, and government.

Improving our defensive network posture is not just about changing equipment and infrastructure; it is also about adopting a proactive defense mind-set. Instead of waiting until an adversary penetrates our networks to assess our vulnerabilities, we have created a specialized team that searches our networks and seeks out those vulnerabilities before they are exploited. This mobile precision capability demonstrates the viability to identify, pursue, and mitigate threats impacting critical links and nodes and provides an additional tool in protecting mission networks. However, we cannot seek or defend everything, so identifying and defending those interfaces that are essential to mission success are crucial. A key facet of this mission is identifying and focusing on a COCOM's prioritized "defended asset list," those critical areas that must be able to operate through an attack. In creating this team, we partnered with the US Transportation Command, as tanker information, logistics tracking, and airlift movements are some of our adversaries' highest-valued targets. As yet a nascent capability, this team may represent one of the most viable missions for expansion.

Proactive defense also reduces the need for human in-the-loop processes; it is far superior to the current reactive process. When we detect an intrusion attempt, the Air Force CERT identifies the characteristics of that attack and updates active sensors, located at multiple defensive levels within the network, with the "learned" information so they can deter existing threats and repel the next attack using the same method. We share information with our academia, industry, and government partners so similar methods of attack can be thwarted across the domain. Our goal is to move away from this reactive process and develop a heuristic capability. Rather than operators having to inform the sensors about each new attack attribute, the sensors themselves will recognize and repel similar attack patterns. Automating this process would further allow us to devote capacity to expanding defensive or mission assurance operations.

Previously, we did things for the sake of the network itself, as if it were the end objective. This resulted in defending every part of the network essentially the same. Our defensive architecture was deployed to defend critical mission systems, core services, and business systems equally. Our primary defensive organization, the Air Force CERT, could not easily distinguish critical mission systems from routine business systems at a base. Today, this is changing. Emphasis is on supporting operational missions

dependent on cyberspace. The focus is on the mission, not the network. This fundamental shift in perspective has driven both how AFSPC crafted the AF Cyber Core Function Master Plan and how AFCYBER refocused its operational activities.

Operationally Leverage Cyberspace Capabilities

Cyberspace operations encompass more than the management and configuration of hardware and software. The Air Force can leverage cyberspace to create integrated effects to respond to crises and conduct uninterrupted operations. As mentioned earlier, instead of responding to the cyberspace environment, we can leverage it to our advantage and our enemies' disadvantage. This provides myriad opportunities to develop and provide new capabilities to the war fighter while offering our adversaries new avenues of attack if we do not fully understand the environment we have created. The repercussions of this new environment must be considered when developing tools and extending the domain to austere locations.

We have come a long way in changing our priority from network assurance to mission assurance. Airmen have begun to distance themselves from a "service provider" maintenance mentality and transition to a "complete the mission" focus. A great example of efforts in this area is support to RPA missions and the objective of operating through a cyberspace attack or outage and accomplishing the mission. Providing mission assurance required extensive front-end mapping to understand the various links from the United States to the overseas flight. The system was designed with over 100 touch points, many of which are not military-controlled, across several different networks, making it critical to establish relationships with commercial organizations. The forward commander of joint air assets prioritizes the most critical RPA missions, and then our operations center identifies and takes proactive steps to ensure the availability of key nodes and failure points along the network infrastructure. While we cannot assure every RPA, we can focus our resources on the highest-priority missions to deliver the greatest downrange advantage. This provides a stark contrast to previous net-focused priorities that resulted in equal defense across the network.

In addition to mission assurance, we are engaged in global operations as the Air Force cyber force provider to US Cyber Command. Over the past two years, our operational units have conducted 17,000 computer net-

work operations in support of combatant command and national agency taskings. Our Airmen executed pursuit of an HVT through computer network exploitation that enabled special operations forces to eliminate the target. We have directly supported objectives to disrupt terrorist command and propaganda efforts. Cyber represents an alternative; it can provide kinetic effects while using nonkinetic capabilities.

COCOMs are beginning to recognize these alternative capabilities and incorporate cyber early in the campaign planning process. Lt Gen Michael Basla, while Air Force Space Command vice-commander, said senior commanders had asked him for the “menu of nonkinetic cyberspace capabilities so they can integrate those into their planning processes.” Cyber capabilities are driving a change in the way we plan, and they require flexibility and a focused, detailed understanding of the cyber environment. We are leveraging the Air Force intelligence community to achieve full-spectrum mission objectives.

To support theater planning for operations in and from cyberspace, target development plays a key role in application of capabilities, especially with respect to industrial control systems (ICS). Rail yards, ports, and power plants are generally built in the same manner worldwide, whether in Tennessee or Ukraine. The initial 80 percent of system understanding can be performed with industry research; the last 20 percent of interface with a particular system requires substantive effort to establish the connections necessary for effective capability employment. Similar to our defensive discussion in figure 1, we currently provide a niche capacity and nascent capability to the war fighter. With constant cyber funding and resources gained from proactive defense, OPLAN-level niche targets, such as ICS infrastructure, offer opportunities to expand combat effectiveness in a resource-constrained environment.

There is a lot of angst on the issue of authorities, and most of it stems from a lack of understanding of how to leverage the necessary authorities to accomplish the mission. Flexibility within the law allows leveraging all the authorities necessary to accomplish the mission without necessarily having a position that bestows the authority on 24 AF. War fighters routinely operate within their inherent Title 10 roles while leveraging the NSA’s SIGINT authorities (Title 50) to support planning and targeting requirements at the tactical, operational, and strategic levels. War-fighter requirements are submitted to the NSA via the national SIGINT requirements process (NSRP) and are vetted and serviced based on national and theater

priorities. This system works well and has been tested in the crucible of war many times. Likewise, 24 AF has units assigned, which are Title 10 units but have a US Signals Intelligence Directive (USSID) that defines the limits and processes they use to collect signals intelligence under the oversight of the Air Force Intelligence, Surveillance, and Reconnaissance Agency and the authority of the NSA. These units routinely move between conducting missions under both their Title 10 and Title 50 hats.

Title 32 authorities define how National Guard units support their respective state. Oft time Air National Guard forces can rapidly transition from Title 32 to support Title 10, all the while exercising caution to ensure Guard members are not put in positions exceeding their authority. For example, when an Air National Guard F-16 is on alert supporting NORTHCOM's air sovereignty mission, it can be training under Title 32, but when it is scrambled, it immediately transitions to a Title 10 role. Conversely, when a natural disaster strikes a state, active duty forces are limited in what they can do under Title 10, but National Guard forces from that state, under the direction of their governor, have more flexibility. This is important when we look at operations in the cyber domain, especially associated with the nation's cyber infrastructure. Industrial control systems are becoming ubiquitous and operate everything from power, water, and fuel systems to building alarms and environmental systems. Title 10 forces assigned to 24 AF have the authority to assess and defend the ICS on a military base. However, they have no authority to deal with systems off base that are essential to military operations. This is a Department of Homeland Security (DHS) responsibility. Though, under certain circumstances, National Guard units, when invited by the civilian entity or acting under the authority of their governor under a declared state of emergency, can be called up to defend of these systems. Interagency policy must continue to evolve and enable these units to synchronize efforts between National Guard and active duty forces to ensure the mission is not interrupted by attacks on the ICS infrastructure off base. Sharing of intelligence and vulnerabilities must also be improved. Today, the national ICS CERT at Idaho National Laboratory performs this function under the authority of the DHS. Synchronizing the ICS CERT efforts with military ICS defensive measures must continue to improve if we are to provide a comprehensive defense of our critical national infrastructure.

Twenty-fourth Air Force can also leverage law enforcement authorities (Title 18) when necessary through our embedded Office of Special In-

vestigations (OSI) support. The OSI works with other law enforcement agencies to investigate cyber crime impacting Air Force networks.

Protecting our information lines of communication and understanding the adversary's key information lines of communication are within the 24 AF's set of responsibilities. We must consider information our key center of gravity and understand what particular information is mission critical to our success. This is not as easy as it may first seem. Are precision navigation and timing our most valuable information, or are timely communications with our airborne assets, including control links to our remotely piloted aircraft? We could expand this list considerably, but the point is made. The difficulty comes when we map the information flows to the supporting infrastructure. Without this level of detail, we cannot adequately defend mission-critical information.

We must also analyze the information centers of gravity of our adversary. This obviously includes those information lines of communication essential to its military operations, but it also includes other information lines of communication that impact the adversary's populace, allies, and supporting entities (including nonstate actors). Similarly, it is critical to understand the information lines of communication that support the adversary's infrastructure, including machine-to-machine communications. By understanding these essential information pathways and systems, we can produce strategic effects without ever staging our forces near an adversary's weapon systems.

Build and Deliver Combat Power

A proper foundation is critical to building a strong structure. It starts with early exposure to science, technology, engineering, and mathematics (STEM). The Air Force supplements the foundation with formal training to create the skilled technical workforce required to manage and protect its cyber resources and facilitate mission users.

A successful STEM program requires collaboration and partnerships with local and national academic and civic leaders. At the high school level, CyberPatriot is the premier national cyber defense competition. It inspires students toward careers in cyber security and other STEM disciplines. At the college level, students compete at the National Collegiate Cyber Defense Competition, and future cyber defenders test their acumen in the National Security Agency's Cyber Defense Exercise. For Reserve

Officer Training Corps cadets, the Advanced Course in Engineering summer program consists of an instructional component and cyber war games, hands-on internships, and cyber officer development that focuses on the study of cyber and its unique leadership challenges. The Air Force Academy's first cyber competition team won the 2012 Cyber Defense Exercise while competing against other service academy cadets, DoD postgraduate students, and the Royal Military College of Canada. In the same week, the team traveled to San Antonio, Texas, and placed second in the National Collegiate Cyber Defense Competition out of 136 teams. In such a dynamic environment, relying only on a STEM background is insufficient for continued success. That is why the AF has established deliberate processes for training and certification of its cyberspace professionals. Undergraduate cyber training (UCT) is a rigorous six-month program to provide foundational training for new cyber officers and enlisted personnel. Intermediate network warfare training builds on UCT and delivers qualified operators prepared to serve in a wide range of positions. Mission qualification training provides unit and position-essential instruction. Similar to the Space 200 and 300 programs, cyber professionals attend Cyber 200 or 300 taught by the Air Force Institute of Technology. These courses provide the career force with continuing education. Last month, we borrowed a page out of our air and space domains by graduating the first weapons instructor course class at the Air Force Warfare Center at Nellis AFB, Nevada. This course teaches professionals to integrate capabilities across air, space, and cyberspace to deliver precise effects. In an effort to increase joint capacity, our sister services are invited to participate in future classes.

DoD training and certification standardization, to include the Guard and Reserve, is key to the nation's success in cyberspace. To emphasize the need for the same training and certifications, the organized Reserve Corps was formally established in 1948 by the Truman administration, but it was not until 1973 when Secretary of Defense James Schlesinger declared the Total Force policy. The Air Force Reserve was held to the same readiness standards and inspections; mobilization planning, operational evaluation, and participation in exercises enhanced Air Reserve Component (ARC) capabilities. In cyber, we can incorporate that same readiness standard, but we must leverage the ARC differently than we have traditionally. We require associations, with flexible drilling, that allow Guard and Reserve members to perform active missions, not merely training scenarios. In the dynamic cyberspace environment, continued engagement is the best way

for the ARC to both support our substantial steady-state mission requirements and be optimally trained and prepared to mobilize, if needed, for a more robust cyber defense of our nation. That continued engagement by our citizen Airmen also enables us to leverage private-sector skills while at the same time providing knowledge gained from bona fide mission experience that should be beneficial to civilian cyber roles in local communities and improve the defenses of industry and government, bringing mainstays of cyber to Main Street. This fuels collaboration between the DoD and the private sector and raises the overall level of national cyber security.

Within the strategy document titled *Sustaining U.S. Global Leadership: Priorities for 21st Century Defense*, Secretary of Defense Leon Panetta makes clear that cyberspace forces are a key component to the nation's ability to project combat power. Specifically, "Modern armed forces cannot conduct high-tempo, effective operations without reliable information and communication networks and assured access to cyberspace and space." To provide resilient, cost-effective cyberspace capabilities for the joint war fighter, an innovative, rapid, tool development process must be accompanied by an acquisition program that reflects an immediate-, medium-, and long-term systems approach.

A factor that hinders the development of cyber capability is the outmoded practices, policies, and rules that guide cyber acquisition from the top down. The current acquisition system was constructed and optimized to support the acquisition of large-scale weapon and training systems. It is based on the five-year Program Objective Memorandum (POM) cycle, which starts two years out from the beginning of the planned acquisition. This drives us to develop large acquisition programs that can survive the vetting process within the Air Force and the Office of the Secretary of Defense. These programs are built from requirements that are defined years in advance and remain relatively static throughout the POM process. The end result is acquisition of outdated equipment and inflexibility that prevents adapting leading-edge technology while it is still leading edge. One example is the modernization of the Air Force boundary. Prior to 2010, the Air Force boundary was defined by more than 140 Internet points of presence, one at each base. But since 2003, we have been consolidating these Internet gateways into 16 regional gateways that now define the boundary to the Air Force network. While the benefit of consolidating the boundaries is indisputable, the "controls" on program execution illustrates the challenge with applying traditional acquisition methodology to cyber

modernization and domain design. Planning for the program began in 2003, and the final gateway was fielded in 2010. By the time the last gateway was fielded, the equipment was obsolete. Although certainly willing to innovate, the process prevented alternatives which kept pace with an intensely dynamic man-made, necessitating modernization of the gateways as soon as they were fielded.

Complicating things further, acquisition programs often field capabilities without a clear understanding of their operational impact on the defensibility, operability, and sustainment of the domain (on behalf of all who use it). Standard acquisition practices often resulted in the fielding of multiple brands and/or standards of network components such as routers and firewalls, adding to the operational burden for the units maintaining and operating the equipment. For example, the Air Force network infrastructure from DISA to the base boundaries includes 1,800 same-brand network routers and switches. Personnel trained on that standard brand are very skilled at operating and configuring those routers. However, a subset of bases deviated with four different brands or variants of routers and switches...without interface testing or a standard for configuration. A small communications team on a base can be trained to efficiently operate nonstandard gear, but as operations are consolidated at network operations units that have enterprise-wide responsibilities, it places undue strain on significantly reduced resources. In theory, these dissimilar infrastructure devices should all communicate with little difficulty, and configuration should be similar. But it does not work that way. While this adds diversity to the network, the ultimate result is a highly heterogeneous network architecture that significantly complicates updating and maintaining these devices. Central management becomes difficult if not impossible, and inevitably, some of the devices do not get properly configured and thereby create vulnerabilities. In addition, training and manpower requirements to maintain such a heterogeneous network cause an unacceptable burden on the already limited cyber manpower resources. This creates a huge workload for Air Force network operations units and adversely impacts the reliability of service to some bases. This problem will be exacerbated as the Air Force continues to offload work from the shrinking base communications units to the network operations units.

One additional innovation involves Air Force Material Command (AFMC) working with AFSPC to establish a Cyber Solutions Center in San Antonio. This center of cyber innovation primarily supports rapid

acquisition providing cutting-edge capabilities for the joint war fighter. It has acquisition professionals from AFMC, science and technology expertise from Air Force Research Laboratory, and is integrated with the cyber development expertise resident in the 24 AF. This team of acquisition, technical, and operational experts is integrated with the daily operations of 24 AF and becomes a powerful engine for innovation that greatly increases the Air Force's ability to create and integrate new and innovative technology. This type of collaboration, along with DoD standardization, increases the capacity of a skilled technical workforce to leverage full-spectrum capabilities to meet the Air Force vision of global reach, global power, and global vigilance.

One opportunity 24 AF is working, in close coordination with AFSPC leadership, is revamping the current program for increasing bandwidth and connectivity at the bases. The legacy program is primarily focused on older, wired technology and fails to leverage the capabilities available with today's wireless technology. By leveraging new technology, we will provide ubiquitous connectivity to base users, reduce infrastructure, increase reliability and resilience, and enable control of government-owned devices to enhance productivity.

Conclusion

Twenty-fourth Air Force is extremely proud of the part its Airmen play in defending the nation in cyberspace at the "speed of cyber," that is, Mach 880,000. The Air Force core contribution to specific joint operations and to the nation's defense is its ability to command, control, and precisely apply forces to provide inherent reach, power, and vigilance—globally. We have effectively leveraged the cyber domain to enhance these core capabilities and to expand operational effectiveness in every engagement. However, this drives a dependency on the networks that directly exchange critical information, often with little human involvement. This trend is only going to increase, as is the trend for adversaries to undermine or contest our ability to leverage the domain. We cannot revert to the days when we, and our platforms, operated without reliable, near-instantaneous access to information—time marches on, and innovators surge forward. 

The Specter of Non-Obvious Warfare

Martin C. Libicki

Innovations, both technological and organizational, over the last few decades have created a potential for non-obvious warfare,¹ in which the identity of the warring side and even the very fact of warfare are completely ambiguous.

The Stuxnet computer worm is only the most recent widely publicized example. This worm is believed to have infiltrated Iran's Natanz centrifuge facility, causing equipment to destroy itself over a period of weeks and leading to the premature retirement of 10 percent of Iran's uranium enrichment capability. Within several months of the worm's public disclosure (September 2010), Western intelligence sources announced that the earliest date Iran could build a bomb had been pushed back several years. Until the worm was discovered and dissected, the Iranians were uncertain why their equipment wore out so fast. Indeed, when confronted publicly with the possibility, they first denied that any such attack had happened, only to reverse themselves obliquely two months later.

Although non-obvious warfare can be epitomized by cyber warfare,² states can attack one another in many ways without the victim being certain exactly who did it or even what was done. Some, like electronic warfare (against nonmilitary targets) and space warfare, have yet to materialize in any strategically significant way. Others, such as naval/land mining or sabotage, have long historical antecedents. What they share is ambiguity. A short list of warfare types that *could* plausibly be conducted in a non-obvious manner includes

- cyber warfare;
- space warfare;
- electronic warfare;

Martin C. Libicki, PhD, is a senior management scientist at RAND Corporation, focusing on the impacts of information technology on domestic and national security. He has published *Conquest in Cyberspace: National Security and Information Warfare* and *Information Technology Standards: Quest for the Common Byte*, as well as numerous monographs. Prior employment includes the National Defense University, the Navy Staff, and the GAO's Energy and Minerals Division. Dr. Libicki holds a master's degree and a PhD from the University of California–Berkeley.

- drone warfare;
- sabotage, special operations, assassins, and mines;
- proxy attacks;
- weapons of mass destruction; and
- intelligence support to combat operations.

Non-obvious warfare stands starkly in contrast to, say, a tank invasion across the German-Polish border, an event unlikely to spur questions such as whose tanks are those . . . and why are they here? By contrast, the *uses* of non-obvious warfare are limited. It is quite difficult to take over the capital of another country anonymously (proxies may do so but at that point often cease being proxies and evolve to dependents or even independents). Defensive warfare is almost always carried out by whomever owns what is being defended. Even coercion requires self-identification *if* the “me” in the point—“don’t tread on me”—is to be adequately conveyed. But there are some types of warfare that can be satisfactorily or even more advantageously carried out if there is doubt about who did what. Again, Stuxnet provides an example. Retarding the Iranian nuclear program benefitted Israel, whether or not anyone knows for certain whether Israel (or anyone else) did it. Furthermore, if the purpose of warfare is to change minds in the victim’s capital, uncertainty may focus subsequent reflection on what such an attack says about the security and (reduced) power of the victim rather than on the malevolence of the undetermined attacker.

Accordingly, this article explores the topic in several steps. The first is to develop a sense of what it means to be non-obvious. The second is to delineate several forms of warfare that may, under some circumstances, be non-obvious and why. The third is to speculate on how states (and non-state actors) might use non-obvious warfare. The fourth is to speculate on how victimized states can respond to the threat of non-obvious warfare.

When is Warfare Non-Obvious?

Ambiguity is the heart of non-obviousness. If the victim is unsure of who carried out an operation, it may hesitate to respond in the same way as if it were certain. Alternatively, the rest of the world might have doubts even if the victim is certain, leaving the victim wary of responding as it might have if *others* were very sure of matters.

Non-obviousness is enhanced if the events in question can themselves be questioned. Some could be accidents or utter mysteries, for example, the unexplained failure of a satellite. Others could be crimes, such as bank robberies by politically inclined groups, or acts of espionage—many events labeled as cyber attacks are really attempts to steal information. Nevertheless, some non-obvious warfare incidents would clearly be acts of war if they were obvious—in which case, the key ambiguity is the actor not the act.

Some forms of warfare are non-obvious because the relationship between the attacker and a state is unclear; for instance, to what extent is Hezbollah working for its own ends, and to what extent is it a puppet manipulated by Tehran? In some cases the perpetrators may be state employees that are not necessarily, or at least not provably, working under the command and control of the state itself. Does the fact that someone close to the Russian political structure claimed credit for having organized attacks on Estonian institutions in Russia mean it was an attack by Russia?³ Pakistan's ISI intelligence agency has been accused of shielding Taliban warlords; so, is Pakistan at war with Afghanistan? If both questions can be answered "yes," then these are two examples of non-obvious warfare.

Finally, many forms of non-obvious warfare present no personal risk to war fighters—which it would have to, almost by definition, since the capture or identification of the perpetrator may make the source of the attack obvious. But one cannot conclude that *states* that employ such war fighters are off the hook just because their war fighters are. A no-fingerprints approach to warfare may be a logical next step after a no-footprints approach, but the two are still quite different.

Non-obviousness is not an absolute, and the actionable response threshold for the victimized state will vary greatly. The primary criterion is how confidently the victim feels a particular state carried out an attack—if, indeed, what happened really *was* an attack. This perceived likelihood is almost always going to be nonzero. Few states truly believe that no other state wants to harm them. Even what later prove to be accidents (e.g., the explosion in the USS *Maine*) is often blamed on other states (e.g., Spain). If there is a crisis (e.g., Spain's attempt to quell a Cuban insurgency), the tendency to believe that any harmful and unusual occurrence was an attack will be that much higher.

So the attacker who would strike with impunity must ask whether or not the confidence with which the victim believes that it carried out the attack

is likely to be greater or less than the confidence that the victim requires to respond to the attack. Everything depends on what the threshold of response is, and there may be many types of responses. Evidence sufficient to gain a criminal conviction in a US court “beyond reasonable doubt” is rarely the issue, although similarly high levels of confidence may, in fact, be required before the victim decides to go to war. On the other hand, mere suspicion may suffice to curtail active or disapprove prospective cooperative arrangements such as mutual military exercises, joint research, or network peering relationships. With some forms of non-obvious warfare, the target may be uncertain of state sponsorship but may convince itself that such a state has to shoulder some blame if it reasonably could have detected and stopped or hindered such an attack and refused to do so.

Exactly how the target state acquires the confidence that another specific state carried out an attack will also vary, but one cannot go very far wrong by considering means, motives, and opportunity. Opportunity—in the form of some traceable delivery vehicle—often best distinguishes obvious from non-obvious warfare. But opportunity is only one leg of the triad. Consider, for example, how the United States would react to the detonation of a so-called suitcase nuclear weapon circa, say, 1962. The suitcase would be incinerated, leaving little forensic evidence. But at that time, only three other states had the *means* to carry out a nuclear attack, and of those three, only one, the USSR, had a *motive* to do so. In such circumstances, the lack of a visible delivery vehicle would have little dented US confidence in the belief that the USSR had done it. Similarly, for many types of non-obvious warfare, such as attacks on spacecraft, the list of suspects would be fairly short since the number of space-faring nations is limited (although, in that case, the victim must also credibly distinguish accidents from attacks).

Types of Non-Obvious Warfare

What makes various forms of non-obvious warfare, in fact, non-obvious? We examine them individually.

Cyber Warfare

Hackers can sit anywhere and attack systems around the world, disrupting their functioning, corrupting the information they hold and the algorithms they run, and, as Stuxnet showed, even breaking machines by

feeding them harmful commands from hacked systems. Attribution is particularly difficult for a cyber attack. The ones and zeroes that constitute the attack do not bear the physical residues of their operators (especially if these ones and zeroes are copied from others' tools). Successfully attacked systems, almost by definition, cannot distinguish an attack from completely benign inputs at the time (with a distributed denial-of-service attack, it is volume, not content, that matters; the attacking bytes generally come from "innocent" machines that have been tricked into spamming the victim). Forensic methods such as tracing the attack back to its sources can be easily frustrated by bouncing the attack through enough portals, using the services of an innocent machine, or jumping on a third-party Wi-Fi connection. Difficulties in attribution may well be inherent to the medium and unlikely to be improved upon in coming years. States wanting to guess who attacked them find they must rely on means and motive. Means offer only a little help for an unsophisticated attack, since over 100 countries have investigated offensive cyber war and the list of hackers includes organized crime groups, nonstate actors, and individuals. It is generally believed that only a state could have pulled off a sophisticated attack such as Stuxnet, with its four zero-day exploits and two stolen certificates. Iran may have figured, once it realized that it *had* been attacked, only Israel and the United States would have both the reason and the talent to carry out such an attack. But it is not entirely impossible that either Russia or China may have wanted to retard Iran's rush to nuclear weapons.

No one yet knows whether cyber attacks carried out in a non-obvious manner will prove advantageous to those who carry them out. It is by no means clear that Russia's (or Russian) attacks on Estonia or Georgia did it that much good. If Israel attacked Iran in cyberspace, what looks like success may be viewed as the beginning of a new set of military operations, or, alternatively, a very special case that no one else can or need duplicate.

Space Warfare

Satellites normally lose capability from time to time in the depths and darkness of space. An attack on a satellite without the attack vehicle being discovered may come close to the perfect crime. States may want to know what happened, but de-orbiting a satellite may not necessarily be something the satellite was designed to do, may be rendered impossible by the nature of the attack, and will require the expenditure of a substantial amount of fuel. Although post-recovery analysis would likely indicate

what happened, it still may not answer who did it. That noted, getting away with “satellite murder” presents difficulties. The United States has the capability to find every sufficiently large ground-based missile launch and tracks space objects supposedly the size of wrenches (the exact details are undoubtedly classified). Because it has a fairly good idea what every satellite is supposed to be doing, those otherwise employed necessarily get noticed, but the advent of microsats, nanosats, and picosats may complicate detection by subtraction in years to come. Ground-based systems might blind satellites, but the satellites have to be looking at whatever it is that is doing the blinding (hence, indicating where the laser is coming from). The number of states that can buy a launch is much larger than the few that can launch objects into space.

Electronic Warfare

As our wired world becomes increasing wireless, the potential for electronic jamming grows apace. Small generic radiating devices surreptitiously emplaced or scattered about can block GPS signals (at least for commercial receivers) and wreak havoc with communications, ranging from cell phone and emergency communications to machine controllers. Such devices can sometimes be quite difficult to find but not hard to characterize (deliberate jamming is unlikely to be confused with natural causes or accidents for very long). Using generic devices can frustrate trace-back, but the real trick in anonymity is to not get caught emplacing such devices. Once the devices start operating, their lifespan is limited, either because they are discovered or because their batteries die.

Drones

Under some relatively narrow set of circumstances, an attack by drones may be carried out without firm attribution. The requirements are many. The drone has to avoid crashing (or must be recovered if it does); otherwise, there is a fair chance of tracing even a generic drone back to its last buyer. The targeted country either has to have relatively poor radar coverage or abut territory or oceans where there is no radar coverage. If the drone comes from the ocean, the list of possible attackers can be limited to those with ships in the area at the time. The drone itself has to be fairly generic—so that its profile at a distance is consistent with the inventory of many different countries—or else stealthy. Finally, the possibility that a drone attack can be a non-obvious attack by the United States must

await the development of attack drones by countries *other than* the United States—failing that, any such drone will be assumed to be American. For states on the outs with the United States, the combination of motive and means may suffice.

Special Operators, Saboteurs, and Assassins

As with drones, the key to maintaining anonymity in special operations is to avoid getting caught. Ironically, the ability to carry out *many* special operations without getting caught requires so much organizational and professional skill that the number of countries capable of doing this is few—making accusations that much more credible. Hence, perfection may be its own undoing, unless the attacker shows considerable restraint. This category includes mine-laying by stealthy conveyance (e.g., submarines), which gives it a historic resonance, if nothing else, but also contemporary resonance, as in the mysterious—and disputed—damage to an Irish vessel primed to run Gaza's blockade.⁴

Proxy Attacks

This broad category includes terrorists, insurgents, militias, and privateers. Attribution becomes difficult because it generally requires the perpetrators be caught (or use a recognizable *modus operandus*) but mostly because it requires tying the perpetrator to a major actor. In practice, however, the link between insurgent groups and states really is ambiguous, and not necessarily by design; empowering individuals with organization, ideology, and weaponry tends to make them believe that their goals are important in and of themselves. The Vietcong, for instance, may have been established and sustained by North Vietnam but had somewhat different priorities.⁵ Africa provides a more apropos case in which various countries that sponsored insurgencies against their neighbors managed to find themselves under siege by insurgents of their own, similarly backed.

Attacks Using Weapons of Mass Destruction

The so-called suitcase bomb of the Cold War era has been joined by the use of biological and chemical agents—of which there are many types—all of which offer, at least in theory, a method of killing people without a state necessarily getting caught doing it. Because weapons of mass destruction, as a general rule, are relatively small, their use may not require forcible insertion, and modern electronics allow them to be detonated

remotely. However, such attacks are considered particularly heinous, and nearly every state has signed one or more international treaties against doing so. For that reason, more such attacks may well be traced to their ultimate source than a similarly stealthy attack by high explosives. Granted, infectious agents, particularly those that may yet be invented by DNA recombination techniques, can be delivered in a very stealthy manner. But unless a state's own citizens are somehow immune to their effects, it is unclear what that state would gain from using them or, if used in a "doomsday machine" mode, why a state would want to be non-obvious about the matter.

Intelligence Support to Combat Operations

Although technically not warfare, a state with a sophisticated stand-off intelligence collection and processing/distribution mechanism can provide data that can be a great help for its friends. If the assistance is not directly intercepted and its distribution is limited, then others would have difficulty discerning the origin for certain (although states may suspect that opponents punching over their weight may have gotten some help, only a handful of countries could and would supply it). Unlike other forms of non-obvious warfare, helping out with information is not particularly heinous, and denials—or at least "neither confirm nor deny"—are par for the course in the intelligence world. Nevertheless the supplying state may not want to show its hand in the conflict lest it be accused of being a belligerent or if it has a rival that can then justify *its own* assistance to the other side.

It merits repetition that unless the attack looks like a complete accident—and the target is completely credulous—there is no such thing as a completely unattributable attack. Every state has its enemies or untrustworthy friends, and if anything untoward happens, the usual suspects will be trotted out for examination. Conversely, plausible deniability matters only if the victimized state really does need something close to judicial proof to take action or is relieved that the authorship of the attack is not so obvious that its unwillingness to respond is not seen as cowardice. Perpetrators do not have to be caught red-handed to suffer reprisal in the hands of those who can put means, motives, and opportunity together to form a sufficiently robust basis for action.

The Uses of Non-Obvious Warfare

It is often easier to state what *cannot* be done with non-obvious warfare. Its inapplicability for conquest and specific coercion has already been noted. Furthermore, any purpose that requires a sustained series of attacks cannot use a non-obvious warfare technique if the probability of ascription for each attack is nonzero and the probability of ascribing one event is at least somewhat independent of the probability of ascribing another. This rules out space warfare, electronic warfare, drones, and special operations. It may also rule out cyber warfare but is less likely to rule out proxy warfare—where attribution has to be inferred rather than discovered—and intelligence support to warfare.

So what *can* be done with non-obvious warfare? One use is general coercion or dissuasion. Instead of signaling, “if you do this we will do that,” the signal is, “if you do this then bad things will happen to you.” Because the act of signaling itself may implicate the attacker, it helps if the signals come from someone else. Others may be willing to help if there are multiple states with a common interest, such as Vietnam, Indonesia, and the Philippines all opposing Chinese bumptiousness in the South China Sea. These others may also be co-religionists or co-ideologues (e.g., “disrespect our religion and bad things happen to you”). The use of non-obvious warfare for compellance is trickier to pull off insofar as it is easier for disparate entities to agree on what can be condemned than to agree on what should be done.

Another fairly obvious use is sabotage, à la Stuxnet, carried out to deny its target some capability. The difficulty is that sabotage is rather pointless unless it takes place on a very large scale or is somehow associated with an operation (if it is a combat operation, the target might assume that the saboteurs work for the combatants). Even if the damage is permanent, states can generally recover. The attack on the Iranian centrifuges made sense because of the strong desire felt by some countries to hobble Iran’s nuclear program and buy time. Another rationale for sabotage is to push a target past a nearby tipping point, even if this tends to be visible only in retrospect. Otherwise, the consequences of carrying out what could be an act of war may outweigh the gains, even if getting caught is unlikely.

An untraceable attack of sufficient magnitude may also weaken the target prefatory to an armed attack or at least so distract the target that it cannot assign the resources, such as sensors, in-place weapons, or management attention, required to foresee and prepare for what turns out to be

an imminent overt attack. Clearly, if an attack does come, the precursor will cease being a non-obvious attack in retrospect (unless the target has multiple eager enemies, each looking for signs of weakness, in which case, what looks obvious may still be wrong). The advantages of starting in a non-obvious mode are twofold. First, if the initial attack were obvious the target might countermove in ways that would make the attack harder to pull off. It may know where to point its defenses, so to speak; it could rally others to pressure the attacker; or it could even counterattack. Second, if the attack falls short of its objectives, the attacker may cancel the overt attack and remain obscure in hopes of eluding punishment.

Correspondingly, a non-obvious attack may be a test to see if the particular technique works, what the target's defenses are, and where improvements should be sought. It would be an expensive test if the target itself should learn something about its vulnerabilities and thereby have cause to work them and evidence on how to do so.

Non-obvious operations can also help win the wars of third parties. Such help can be non-obvious either if the *fact* of help is not obvious or if the *source* of help could be any of several countries or entities such as insurgent or mercenary groups. This raises the question of why such a state would want not to leave fingerprints. One reason is that the attacks take place in a country other than the one that wanted help (e.g., Syria attacks Iraq, and the United States attacks targets in Syria), thereby becoming an act of war in its own right and an excuse for the attacked country to call on *its* friends to help (e.g., attack Iraq). More likely, however, the assistance supports operations within the state under attack, either by another state or by insurgents, so these factors do not come into play. What *does* matter, however, is the appearance of commitment and how it prevents assuming a commitment to pursue victory or lose face. Intervening and then withdrawing prematurely raises doubts about the state's seriousness of purpose and even trustworthiness, even if such a state never made an explicit commitment to stay the course.

Non-obvious warfare can also be carried out for narrative effect. Normally, in warfare the attacker and the target are both part of the narrative, and unless the attacker's actions are totally baseless, the contest over narratives is likely to be two-handed with each side's fans supporting their own side. However, if the attacker is unknown, or at least unclear, then the focus of the story is necessarily on the target, and the theme is likely to focus on why the target was attacked—and may well dwell on what the target did

that merited the attack or why the target could not secure itself. That, in fact, may be the attacker's motive: to create a crisis of confidence in the target state, either weakening it outright, creating fissures in its body politic, or at least making it more amenable to concession.

Finally, if an attacker can persuade the target that it was hit by a third party, it may catalyze conflict that will be to the attacker's advantage. A non-obvious Taiwanese cyber attack on the United States during a crisis with China, for instance, might put the United States at odds with China and thus more likely to support Taiwan. An attacker that instigates a war between two former trading partners could force both to purchase from the remaining relevant neutral, the attacker. Of course, if attribution follows, the attacker will have made one enemy it did not need and perhaps a second enemy as well—the country that the attacker hoped would be fingered.

The Target's Response Options

In some cases, ambiguity works to the target's advantage by giving it an excuse to avoid responding; it can claim uncertainty about who perpetrated the attack or what, in fact, was done. Not knowing helps the targeted nation ward off popular calls to fight and redeem its honor. In some cases the attacker itself may not necessarily think the worse of the target's honor if no response ensues; in other cases, it will convince itself the target knew but was lying to avoid a confrontation. Consider, analogously, the phantom Israeli nuclear arsenal. Once other powerful Middle Eastern states acknowledge that Israel has nuclear arms, they must answer as to why they do not. No polity is fooled, but neither must it be taunted by the prospect.

Mostly, though, targets would simply want such attacks to stop—but how? Defense is clearly an option and one that would logically assume greater importance the less it can lean on not hitting back because it is unsure about who committed the offense. Another option is to help create pressure from the world community to end the possession of the requisite attack technology, but most of these cannot be effectively banned. Cyber weapons are largely the obverse of system vulnerabilities, the attack code is trivial to hide, and the underlying technologies of offense are required for cyber defense. Electronic jamming is inherent in the ability to generate radio frequency energy. Intelligence support for third parties is identical

to intelligence support for military operations in general. The weapons of sabotage, special operations, and insurgencies are small arms. Conversely, weapons of mass destruction and land mines (but not naval mines) are already banned by treaty. The only weapons not covered by treaties that could conceivably be banned are antisatellite weapons and drones; both have legitimate (overt) military purposes. More broadly, it is how such weapons are used rather than the weapons themselves that determines the characteristics of non-obvious warfare.

A variant on the second approach is to develop a global consensus that the covert use of warfare is far more heinous than its overt use. Thus, if such weapons *are* used—something that may not always be apparent—the world community would support efforts to pressure potential users into allowing investigations that would clarify which state was at fault. After all, most forms of warfare are universally held to be crimes if carried out by those outside the military; thus, even the accused state should have an interest in finding and rooting out its dangerous criminals, assuming it would wish to shift the blame. Where states use proxies and such acts *are* crimes, they may be pressured to cooperate with international police investigations. Satisfaction for the aggrieved party, however, assumes police actions can establish reasonable levels of certainty. More problematically, the closer the trail of investigation comes to the doors of military or intelligence establishments, the greater the reluctance of states to allow matters to proceed. Such reluctance would not be unfounded—if purported acts of non-obvious warfare allow investigators to peer into covert operations, states may go to great lengths to interpret the need for evidence in ways that would also allow them to uncover the secrets of their rivals.

The last recourse is for victimized states and their allies to respond to suspected warring states as if certain they did it. In doing so, they must factor in how certain *others* are that the accusation is correct and, to some extent, whether the purported attacking state believes it is guilty. Many non-obvious warfare techniques can be carried out by rogue elements. As noted, some responses, such as chilling relations between the target and the purported attacker, do not require anything close to conclusive proof; mere uneasiness suffices. Other responses, such as retaliation, normally require high levels of confidence. In the end, the victimized state has to weigh the risks associated with false negatives (doing nothing in the face of aggression) and false positives (retaliating against the innocent). Note further that “plausible deniability” is hardly an absolute in this case. Unless the

victimized state can only respond through the court system—and states cannot go on trial, only their leaders—the balance between responding and not responding may tip well before the confidence meter hits 100 percent. A relatively pacifist state surrounded on all sides by friends (e.g., Belgium) and embraced by alliances may want near certainty and may not react even then; an anxious, well-armed state surrounded on all sides by potential adversaries (e.g., Israel) may be less fussy.

Or the victim could retaliate by using non-obvious warfare itself. Ostensibly, the mutual commitment of both sides to modulate their responses to one another might limit the potential for open and, hence, more destructive warfare—as long as both sides are careful not to reveal themselves. This may create a set of strange incentives wherein both sides' non-obvious warfare communities take pains not to reveal the activities of their counterparts lest power and influence on both sides shift to communities whose warfare methods are quite obvious. Conversely, the perception that it is acceptable to escalate in a non-obvious manner rather than call out the other side may allow the destructive cost of non-obvious warfare to rise to its limits. If matters then become obvious, the warfare level that forms the foundation for the next set of threats starts at the much higher level.

Assessment and Conclusions

Would the spread of non-obvious warfare be a good thing? Even if wielded solely in pursuit of good aims, such techniques corrode both military values and diplomatic norms. Non-obvious warfare, almost by definition, has to be the work of small teams that must isolate themselves from the larger community, much like intelligence operatives, lest word of their adventures leak out. The efforts of the small non-obvious warfare teams would leave the mass of the national security establishment quite uncertain about what exactly was going on and who exactly was behind all the activity (only some of which would appear to be accidental).

Non-obvious warfare is also a poor fit for democratic states and a far better fit for authoritarian or failing states in which the intelligence community has become decoupled from its legitimate governance structure. States with long-term reputations to manage are likely to see the downside from having to lie about their warfare activities when so confronted.

Universal or even wide adoption of non-obvious warfare would likely yield a more suspicious world. Once attacks are shaped to look like accidents,

many accidents will start to smell like attacks. Nations would react (even more than they do now) to suspicions rather than actual substance; attackers might be credited/blamed for far more than they actually merit. In too many countries, *anything* that seems askew is blamed on the United States (or Israel) and their ubiquitous and omnipotent intelligence agencies. Part of their polities' maturity entails improvements in their ability to distinguish fact from fantasy; evidence that such fantasy had a kernel of truth behind it would hardly facilitate the maturation process. Indeed, under crisis circumstances, it is conceivable a conflict could start even though the accused did nothing. And of course, a crisis could start when a state used such techniques thinking it would never be caught—and was.

Notes

1. The term *non-obvious* had an earlier manifestation in Jeff Jonas's data-mining product, Non-Obvious Relationship Analysis.

2. *Warfare*, used here, comprises operations carried out for political ends by states aimed at the destruction, corruption, or significant disruption of assets or interests associated with other states using means that are generally considered illegal if not done by states. Our discussion is limited to states, because nonstate actors do not always have return addresses or even always unambiguous identities, and individuals therein can be subject to legal actions in ways that states cannot be.

3. Sergei Markov, a state Duma deputy from the pro-Kremlin Unified Russia Party, claimed, "About the cyberattack on Estonia . . . don't worry, that attack was carried out by my assistant. I won't tell you his name, because then he might not be able to get visas." "Behind the Estonia Cyberattacks," *Radio Free Europe/Radio Liberty*, 6 March 2009, http://www.rferl.org/content/Behind_The_Estonia_Cyberattacks/1505613.html.

4. Robert Mackey, "Irish Flotilla Activists Show Damage to their Boat," *The Lede: Blogging the News*, 1 July 2011, <http://thelede.blogs.nytimes.com/2011/07/01/what-flotilla-activists-videos-look-like/>.

5. Which came to near naught after the original ranks were greatly reduced in the 1968 Tet offensive.

Internet Governance and National Security

Panayotis A. Yannakogeorgos

The debate over network protocols illustrates how standards can be politics by other means.

—Janet Abbate, *Inventing the Internet* (1999)

The organizing ethos of the Internet founders was that of a boundless space enabling everyone to connect with everything, everywhere. This governing principle did not reflect laws or national borders. Indeed, everyone was equal. A brave new world emerged where the meek are powerful enough to challenge the strong. Perhaps the best articulation of these sentiments is found in “A Declaration of Independence of Cyberspace.” Addressing world governments and corporations online, John Perry Barlow proclaimed, “Your legal concepts of property, expression, identity, movement, and context do not apply to us. They are all based on matter, and there is no matter here.”¹ Romanticized anarchic visions of the Internet came to be synonymized with cyberspace writ large. The dynamics of stakeholders involved with the inputs and processes that govern this global telecommunications experiment were not taken into account by the utopian vision that came to frame the policy questions of the early twenty-first century. Juxtapose this view with that of some Internet stakeholders who view the project as a “rational regime of access and flow of information, acknowledging that the network is not some renewable natural resource but a man-made structure that exists only owing to decades of infrastructure building at great cost to great companies, entities that believe they ultimately are entitled to a say.”²

Dr. Pano Yannakogeorgos is a research professor of cyber policy and global affairs at the Air Force Research Institute of Air University. His research interests include the intersection of cyberspace and global security, cyber norms, cyber arms control, violent nonstate actors, and Balkan and Eastern Mediterranean studies. He formerly held appointments as senior program coordinator at the Rutgers University Division of Global Affairs and was an adviser to the UN Security Council. He holds PhD and MS degrees in global affairs from Rutgers University and an ALB degree in philosophy from Harvard University.

The sole purpose of cyberspace is to create effects in the real world, and the US high-tech sector leads the world in innovating and developing hardware, software, and content services.³ American companies provide technologies that allow more and better digital information to flow across borders, thereby enhancing socioeconomic development worldwide. When markets and Internet connections are open, America's information technology (IT) companies shape the world and prosper. Leveraging the benefits of the Internet cannot occur, however, if confidence in networked digital information and communications technologies is lacking. In cyberspace, security is the cornerstone of the confidence that leads to openness and prosperity. While the most potent manifestation of cyberspace, the Internet, works seamlessly, the protocols and standards that allow computers to interoperate are what have permitted this technological wonder to catalyze innovation and prosperity globally. The power of the current Internet governance model strengthens the global power of the American example and facilitates democratization and development abroad by permitting the free flow of information to create economic growth and global innovation.⁴ Today, this Internet is at risk from infrastructure and protocol design, development, and standardization by corporate entities of nondemocratic states.

Cyber security discussions largely focus on the conflict created by headline-grabbing exploits of ad hoc hacker networks or nation-state-inspired corporate espionage.⁵ Malicious actors add to the conflict and are indeed exploiting vulnerabilities in information systems. But there is a different side of cyber conflict that presents a perhaps graver national security challenge: that is the "friendly" side of cyber conquest, as Martin Libicki once termed it.⁶ The friendly side of cyber conquest of the Internet entails dominance of the technical and public policy issues that govern how the Internet operates. Current US cyber security strategies do not adequately address the increasing activity of authoritarian states and their corporations within the technical bodies responsible for developing the protocols and standards on which current and next-generation digital networks function.

Internet governance can be defined as a wide field including infrastructure, standardization, legal, sociocultural, economic, and development issues. But the issues related to governance of critical Internet resources and their impact on US national security are often overlooked. Foreign efforts to alter the technical management of the Internet and the design of technical standards may undermine US national interests in the

long term. This article discusses the US national security policy context and presents the concept of friendly conquest and the multistakeholder format of Internet governance which allows for the free flow of information. There are many global challenges to the status quo, including the rise of alternative computer networks in cyberspace, that beg for recommendations to address those challenges.

Internet Governance and US National Cyber Strategy

Technical standards and protocols do not elicit the same attention as more visible threats to national cyber security. In a human capital and resource-constrained environment, attention has focused on crime, espionage, and other forms of cyber conflict rather than on the issues related to governance of critical Internet resources, development of technical standards, and design of new telecommunications equipment. In a domain that is already confusing to policy wonks, the complexity of Internet governance makes it even harder for policymakers to commit resources to a field that has no analogy in the physical world. In the nuclear age, there was no debate as to whether one could redesign the physical properties of uranium and apply them universally to eliminate the element's potential for weaponization. The underlying language of nuclear conflict was constrained by the laws of physics (e.g., nuclear fission, gravity). Physical limits in cyberspace exist as well by constraining information flows to the laws of physics—the wave-particle duality of radiation which, when modulated with bits, creates an information flow. However, the “logic” elements of cyber that permit information to flow across networks and appear within applications to create effects in the real world are bound only by the limits of human innovation. This affects the character of cyberspace. Its current form is free and open, but that does not necessarily mean it always will be. Understanding the strategic-level issues of Internet governance are thus just as critical as understanding the impact of vulnerabilities that attackers may exploit to cause incidents of national security concern. In the national security context, the technical management of the Internet matters because it may allow authoritarian states to exert power and influence over the underlying infrastructure. In the global security context, maintaining the values of free-flowing information within Internet governance bodies will continue to foster innovation and economic prosperity in both developed and developing states.

Several current national strategies articulate nationwide responses to cyber threats.⁷ They tend to focus on catastrophic national security incidents rather than on the battles within the organizations that set technical standards or manage the day-to-day operation of the Internet. The White House does highlight the importance of current multistakeholder forums for design and standardization of the technical standards via “collaborative development of consensus-based international standards for information and communication technology . . . a key part of preserving openness and interoperability, growing our digital economies, and moving our societies forward.”⁸ Furthermore, the challenges we face in international standards-setting bodies are recognized in that “in designing the next generation of these systems, we must advance the common interest by supporting the soundest technical standards and governance structures, rather than those that will simply enhance national prestige or political control.”⁹ However, these issues are drowned out by more-sensational, hypothetical situations of a cyber doomsday.

Security demands that the language of the Internet—the underlying technical standards and protocols—continue to sustain free-flowing information. If “code is law” in cyberspace, as some posit,¹⁰ then the standards and protocols are the fabric of cyber reality that give code meaning. In policy circles, cyberspace is already considered the “invisible domain.” Technical standards and protocols are thus, “invisible” squared. However, these protocols define the character of the Internet and its underlying critical infrastructures. As noted elsewhere, “The underlying protocols to which software and hardware design conforms represent a more embedded and more invisible form of level architecture to constrain behavior, establish public policy. . . . [I]n this sense protocols have political agency—not a disembodied agency but one derived from protocol designers and implementers.”¹¹ In the past it was the United States that led the world in the development of protocols and standards. As a result, the values of freedom were embedded in the Internet’s design and character, which incubated innovation that continues to spur socioeconomic development globally.

Within the DoD context, a single, connected, open Internet is critical to assuring its missions by facilitating collaboration within the agency and with its mission partners. Today, the department lists in its *Strategy for Operating in Cyberspace* its concerns about “external threat actors, insider threats, supply chain vulnerabilities, and threats to DoD’s operational ability.”¹² Other elements from the DoD’s *Information Enterprise Strategic*

Plan that articulate concerns with Internet governance and advocate for “DoD equities at international technical and governance meetings” should be added to the list.¹³ However, the sheer political nature of the documents does not adequately address broader US foreign policy goals within global Internet governance bodies as much as intended. Thus, DoD computer scientists and engineers risk taking the backseat in an area where they once pioneered. Creating the Internet and maintaining the technical edge are two very different problems.

The Friendly Side of Cyber Conflict

Looming battles in Internet standards and governance bodies will determine the future character of the Internet. The advanced deployment of IPv6 in Russia and China and development of new standards by near-peer-competitor countries are creating new technical standards and deploying them into the global marketplace, thus enabling friendly cyber conflict.

Friendly conquest occurs when a noncore operator of a system enters into partnership with a core operator in exchange for access to a desired information system. Cyber theorist Martin Libicki notes,

One who controls a system may let others access it so that they may enjoy its content, services and connections. With time, if such access is useful . . . users may find themselves not only growing dependent on it, but [also] deepening their dependence on it by adopting standards and protocols for their own systems and making investments in order to better use the content, services or connections they enjoy.¹⁴

The core partner in such a coalition emerges to dominate noncore members who have come to depend on the service offered, though not without some vulnerability to the core partner’s network. Fears exist “that the full dependence that pervades one’s internal systems may leave one open for manipulation. . . . The source of such vulnerability could range from one partner’s general knowledge of how the infrastructure is secure, to privileged access to the infrastructure that can permit an attack to be bootstrapped more easily.”¹⁵

Libicki operates with relational mechanisms to explain how coalitions leading to friendly conquest occur. Friendly conquest in cyberspace can be surmised as the willing participation of X in Y’s information system. X willingly enters into a coalition with Y in cyberspace. Y’s friendly

conquest of X occurs when X becomes dependent on Y's system. This is not to say that X merely entering the coalition will cause the conquest. X's perceived need for access to Y's cyberspace (or inability to construct its own) causes it to willingly enter into a coalition with Y. X adopts Y's standards and protocols making up the information system architecture of Y's cyberspace in a way that allows it to interoperate within X's cyberspace. X adopts Y's cyberspace architecture and thus the necessary condition for Y's friendly conquest. It is a facilitating condition for X's hostile conquest. X might begin to use the standards and protocols of Y's cyberspace as a model for its own cyberspace. Since Y is an expert in its own standards and protocols, X's modeling of these standards in its own systems is another vulnerability, which can facilitate X's hostile conquest by Y. X does not have to be a friend. It can be a neutral or a possible future enemy of Y. There is utility in Y opening its cyberspace to X only if Y sees some benefit to itself, although Libicki does argue that Y will open its cyberspace regardless. Once friendly conquest is accomplished, Libicki argues, it can facilitate hostile conquest in cyberspace. Friendly conquest of X by Y may thus facilitate hostile conquest in cyberspace conducted by Y against X.

The Internet and its underlying technical infrastructure is a potent manifestation of how the United States, as core operator of an information system, extended friendly dominance over allies and adversaries alike through creation of the technology and setting the rules for its operation. The Internet relies on products designed and operated by US-based entities such as the Domain Name System (DNS) and Internet Corporation for Assigned Names and Numbers (ICANN), Microsoft, and Cisco. Users around the world, such as Google and Facebook, have come to rely on services offered over this platform. The dominant position that US-based entities currently have is not permanent. The Estonian-developed Skype is indicative that services may be non-US in origin. Yet, even when an Internet-based service is created by foreign entities, most of the information flowing through the said application passes through hardware in the United States. When vulnerabilities are perceived, other nations may try to exit our information system to preserve their cyber sovereignty and expand their influence by attracting customers toward their own indigenous systems and away from the Internet.¹⁶ Thus, our strategic advantage in cyberspace is not timeless and is being contested in varying degrees by near-peer competitors. Hence, we should understand their current

responses to US technological dominance to refine our cyber strategy within the context of friendly cyber conquest.

US Air Force doctrine recognizes one aspect of friendly conquest: supply-side infrastructure vulnerabilities. “Many of the COTS [commercial off the shelf] technologies (hardware and software) the Air Force purchases are developed, manufactured, or have components manufactured by foreign countries. These manufacturers, vendors, service providers, and developers can be influenced by adversaries to provide altered products that have built-in vulnerabilities, such as modified chips.”¹⁷ Friendly conquest goes beyond adversaries merely being able to infiltrate the supply chain and create backdoors on servers of national security significance before they enter the United States.¹⁸ The threat also comes from the emergence of new technologies in which the United States is not the core operator but may become dependent. With the focus on malicious cyber attacks, not enough attention is being paid to the soft underbelly of the cyber world—the technologies and standards that have allowed cyberspace to emerge from the electromagnetic spectrum.

China is making a great leap forward in terms of sowing the seeds for global friendly conquest in cyberspace. As reported by the US-China Economic and Security Review Commission, “If current trends continue, China (combined with proxy interests) will effectively become the principal market driver in many sectors, including telecom, on the basis of consumption, production, and innovation.”¹⁹ US reliance on China as a manufacturer of computer chips and other information and communications technology (ICT) hardware has allowed viruses and backdoors in equipment used by US-based entities, including the military. Extraordinarily low-priced Chinese-made computer hardware is a lucrative buy in Asia and the developing world.²⁰ Furthermore, Chinese entities, such as Huawei, are on the leading edge of developing the standards of next-generation mobile 4G LTE networks.²¹

One example of how efforts at friendly conquest can backfire and make the United States vulnerable to cyber attack was demonstrated in Microsoft’s experience with China. In 2003, China received access to the source code for Microsoft Windows in a partnership with Microsoft to cooperate on the discovery and resolution of Windows security issues. The China Information Technology Security Certification Center (CNITSEC) Source Code Review Lab, described as “the only national certification center in China to adopt the international GB/T 18336, the ISO 15408 standard

to test, evaluate and certify information security products, systems and Web services,” was the focal point of this collaboration.²² Undeterred by International Organization of Standards (ISO) criteria, and unanticipated by many experts in the field, Chinese computer scientists reverse-engineered the code. This allowed them to develop malicious code, including viruses, Trojan horses, and backdoors, that exploited software vulnerabilities in the operating system. These efforts resulted in the shutting down of the US Pacific Command Headquarters after a Chinese-based attack.²³ Chinese entities are also making great strides in developing core information systems upon which others will come to rely. Virtual reality (VR) technologies are one example of an emerging tool that could become as ubiquitous for social and commercial interactions as the Internet is today. Globally, people are increasingly using VR technology fused with the Internet to socially interact.²⁴ Experts have noted that

any country that succeeds in dominating the VR market may also set the technical standards for the rest of the world, and may also own and operate the VR servers that give them unique access to information about future global financial transactions, transportation, shipping, and business communications that may rely on virtual worlds. . . .

Global commerce is expected to “come to rely heavily on VR.” Banking, transportation control, communications are all types of global commerce occurring in a virtual reality.²⁵

While current strategies do address the supply-chain risks posed by foreign manufacturing, the trend of China taking the lead in the protocols that will come to underlie VR and other technologies, as well as standard setting within international bodies, is a challenge that current cyber strategies insufficiently address. This may be due in part to the cultural differences in the relations between US-headquartered multinational corporations (MNC) and the US government (USG) versus the MNCs in foreign countries that at times have very close relations to their own governments.

Multistakeholders and Internet Governance

Business entities such as multinational corporations contribute to the formation of policies regulating international communications formally within the International Telecommunications Union (ITU) and informally through the personal contributions of their employees within the ICANN, the Internet Engineering Task Force (IETF), and other organizations.

Within the United States, telecommunications service providers (dating back to the era of electrical telegraph systems) were never part of a state-owned monopoly. This was not the case in the rest of the world.²⁶ British Telecom and Deutsche Telekom, for example, were state-owned entities before being privatized in the 1990s. Granted, although there is no direct state control within the United States, telecommunications companies are regulated by the state. In international telecommunications negotiations, a state and its ICT firms have a symbiotic relationship.²⁷ This has been the case since the International Telegraph Union, predecessor of the International Telecommunications Union, began meeting in the mid nineteenth century to regulate telegraphy policies.²⁸ Thus, the view in the developing world is that “at present, it is . . . U.S. law which applies globally by default as most monopoly Internet companies are U.S.-based.”²⁹

If trade is a political activity, then firms are political actors. States can utilize firms to distribute or reward power to meet their own political objectives.³⁰ Since states and firms both cause effects on the behavior of the other, a dynamic bidirectional interaction exists between the state and the MNC.

Important policy tools that affect the behavior of MNCs include export controls, protectionism, and strategic trade policy. Export controls tend to have a political purpose since, as one expert notes, “they are designed to prevent rival states from gaining access to key resources and technologies,” or to punish a state.³¹ Firms manufacturing strategic goods rely on governments to adopt trade policies that will support the firm’s competitive stance in the global market,³² but states do place restrictions on what may be exported, even if it is to the detriment of a firm’s competitiveness in foreign markets.³³ In the United States, the federal government lost the so-called encryption wars of the 1990s, when private industry protested policies prohibiting the export of strong encryption software for strategic reasons.³⁴

In an effort to prevent criminals from communicating using unbreakable codes, some firms implement law enforcement intercept (LEI) mechanisms so national security agencies can monitor suspected criminal and terrorist communications.³⁵ US firms and persons associated with them, who develop, maintain, and revise the core standards and technological infrastructures, are stigmatized by such allegations which depict a rogue national security apparatus and private sector in collusion capturing all of the world’s data. This does not reflect the fact that, unlike in authoritarian states, careful compliance with US laws designed to protect user privacy maintains a separation between government and the private sector.³⁶ Media

preferring headline-grabbing allegations decrease global trust in the American private sector and validate the narratives that the Internet governance mechanisms must be internationalized. Thus, the close relationship between governments and firms in the area of strategic trade policy affects both how firms operate and how governments counteract the misuse of cyberspace.³⁷

The global perception that the US government has de facto control of critical Internet resources is largely shaped by other nations' experiences of the close relationship between telecommunications companies and their national governments. Uniquely, the US government has never owned or operated any telecommunications companies. As the rest of the world shifted to the US privatized telecom model, prior experience of government control of the sector did not leave their cognitive balance. Today these experiences cast a shadow of suspicion over the special agreement between the ICANN and the US Department of Commerce.

Critical Internet Resources and Infrastructure

Technical management of the Domain Name System, invented by the DoD and governed by it in its formative years, was assumed by the Department of Commerce in 1998 and subsequently evolved into its current nongovernmental multistakeholder model.³⁸ The description here will not delve into the tactical- and operational-level functioning of each organization that has a role in Internet governance.³⁹ It will instead offer a brief recap of the underlying technology and the organizations that have a role in setting the standards which allow for technical functioning of the Internet. It is thus the purpose of this section to provide an account of Internet governance as a source of national security concern. With discussions focusing on malicious activities, there has been little consideration to the implications of the peaceful work of designing and maintaining the Internet and the implications these activities have on US interests.

Critical Internet resources (CIR) "in the context of Internet governance usually refers to Internet unique logical resources rather than physical infrastructural components or virtual resources not exclusive to the Internet. CIRs must provide a technical requirement of global uniqueness requiring some central coordination: Internet address, DNS, Autonomous System Numbers."⁴⁰ Unlike the popular conception of a limitless Internet, the underlying address space is limited. Indeed, IP address space has nearly run out. Foreseeing this Internet protocol, engineers developed IPv6, which among other improvements increased the total number of potential

IP addresses from 4,294,967,296 in IPv4 to 2^{128} in IPv6. It is recognized today that “deploying IPv6 is the only perennial way to ease pressure on the public IPv4 address pool.”⁴¹ As the world begins a transition from using IPv4 to IPv6 as the dominant communications protocol for the global Internet, the United States is not leading its deployment. Russia currently enjoys the greatest deployment in terms of market penetration, and China enjoys the greatest deployment in sheer numbers.⁴² The consequences of delayed deployment are related to both Internet governance and the more traditional security threats. On the latter point the National Institute for Standards notes that the “prevention of unauthorized access to IPv6 networks will likely be more difficult in the early years of IPv6 deployments.”⁴³ Thus, competitor nations that have more experience in national-level deployments of IPv6 have greater technical understanding of its real-world operations. The Air Force NIPRNet will not be entirely enabled for IPv6 until 2014. Even then, it has been noted that the plan is to use both IPv4 and IPv6 in parallel for the next 10–15 years.⁴⁴ As deployment of IPv6 as the backbone of the Internet continues, Russia and China may have the perceived legitimacy as IPv6 leads and take advantage of that opportunity to shift control of these scarce address spaces from the ICANN toward the control of an intergovernmental body, such as the United Nations.

The ICANN and the Current Internet Governance Structure

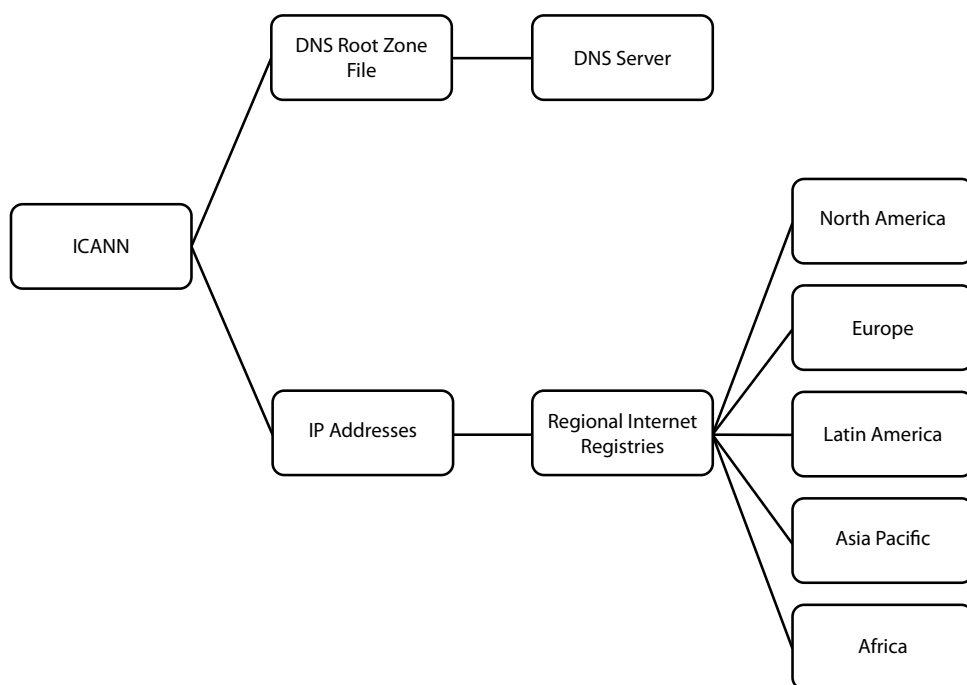
Because cyberspace is a man-made domain, infrastructure and standardization are critically important. Global bodies of computer scientists and engineers create the standards and rules on which the Internet—the most potent manifestation of cyberspace—operates. Indeed, many of these global bodies began as DISA, DARPA, or other USG programs that were privatized in the mid 1990s. Thus, the development of the next-generation Internet does not have the United States as the prime mover.⁴⁵ Instead, standards and processes are being developed by Russian, Chinese, and other foreign scientists and engineers. Today’s machines speak a form of the English language to each other. If US scientific excellence continues its degenerative path, future networks may come to rely on machines speaking foreign languages. Furthermore, governance of the DNS and IP address allocation is being challenged to migrate from the current multistakeholder approach to an intergovernmental mechanism within the ITU. This is the friendly side of cyber conflict.

The DNS allows people to use Uniform Resource Locators (URL) to communicate with other machines on the Internet. Instead of having to type in the IP address of a website—a string of numbers—a person can type a natural language URL, such as *www.af.mil*, into a web browser to connect with the desired corresponding IP address. This makes the web user-friendly, and to the common user, might as well be the work of a wizard that allows information to be piped onto someone's computer. However, IP addresses are scarce, especially in IPv4. The processes for assigning scarce IP addresses and allowing the Internet to serve as a global platform are complex, both technically and, increasingly, politically.

The allocation of IPv4 address space to various registries is provided by ICANN via the Internet Assigned Numbers Authority (IANA).⁴⁶ Globally routable IP addresses reside in DNS databases on root zone databases that allow for the translation of URLs into IP addresses.⁴⁷ (see figure next page). The top-level domain names, such as *.com* or *.org*, are maintained and updated by the ICANN, which was once under the Department of Commerce (DoC). Now operating under a memorandum of understanding with the DoC, the ICANN continues to be the sole source of IP address allocation to specific DNSs and regional Internet registries to assure a uniform Internet experience for all. By governing and maintaining the DNS central root zone databases and backing them up on DNS servers worldwide, the ICANN assures that if a domain name is available, someone can buy it and link it with an IP address to create an online presence.⁴⁸

Internet Engineering Task Force: Stewards of TCP/IP

Internationally standardized communications protocol stack, called Transmission Control Protocol and Internet Protocol (TCP/IP), allows for the flow of data packets and information across computer networks, including the Internet. TCP/IP is standardized by the International Organization of Standards for the Open Systems Interconnection (OSI) model as the basis of Internet networking. A brief description of how information is sent across networks is necessary to better understand the significance of TCP/IP. Data packets are the basic units of network traffic. They are the standard means of dividing information into smaller units when sending it over a network. A significant component of computer networks is the IP header, which contains information pertaining to the source and destination addresses. Machines require these strings of numbers to connect with other computers on the Internet or other networks.⁴⁹ All



networked hardware must have a valid IP address to function on a network. Data packets are recreated by the receiving machine based on information within a header of each packet that tells the receiving computer how to recreate information from the packet data. Without internationally standardized protocols such as TCP/IP, there would be no assurance that packets could be read by a receiving machine.⁵⁰

The most esoteric of all critical Internet resources are the autonomous system numbers (ASN). These numbers are used by network providers at “peering points” to allow information to flow from, say, Verizon to ATT, among other uses. Border gateway protocols are one aspect of ASNs.

Internet policy debates have proven the ineffectualness of multilateralism as the United States strives to lead and others fail to follow. American technological innovation in the development and maintenance of the Internet’s backbone is unquestioned. But global efforts to promote regulatory reform, such as including institutions of global governance like the ITU as entities responsible for overseeing the ICANN, are a tense political issue closely linked with the national cyber security concerns of democratic and autocratic regimes alike. In sum, American “leadership” as first among equals has led to a succession of dead ends. We are witnessing

countermoves by friends and competitors alike that may gain momentum during the 2012 World Conference on International Telecommunication.⁵¹

Global Challenges to the Status Quo

Global information flowing through open elements of cyberspace, such as the Internet, is regulated by national and regional bodies coordinating their policies internationally. Standards that have been created for elements of cyberspace have required lengthy processes at various bodies, such as the International Organization for Standardization and ITU, to assure sufficient technical and political cooperation among nation-states. While US-based entities have traditionally set the standards for Internet technology, China-based entities, such as the ZTE Corporation, are increasingly taking on roles within the ITU to draft important international standards that will shape the world's next-generation networks. This is not a recent development. As early as 2004, Chinese personnel working in senior ITU Telecommunication Standardization Sector positions began to discuss using the transition to IPv6 as a way to correct a perceived imbalance in address allocation between the United States and the developing world: "The early allocation of IPv4 addresses resulted in geographic imbalances and an excessive possession of the address space by early adopters. This situation was recognized and addressed by the Regional Internet Registries (RIR). . . . Some developing countries have raised issues regarding IP address allocation. It is important to ensure that similar concerns do not arise with respect to IPv6."⁵² This is indicative of a desire by some states to perhaps shift the governance of IPv6 address allocation into a global institution such as the ITU.

From the perspective of maintaining US national interests, the current multistakeholder framework governing critical Internet resources continues to be a good mechanism for regulating the day-to-day technical operations of the Internet. However, momentum related to Internet governance within the United Nations is gaining within political forums. Led by Russian and Chinese initiatives, competitors and partners alike have been working toward internationalizing the Internet's technical governance. China and Russia, along with India, South Africa, and Brazil, have led initiatives against US dominance of the ICANN. These efforts have been in the works for nearly a decade.⁵³ As the DoD ARPANET experiment emerged to become a significant component of global socioeconomic development and governments increasingly came to realize its importance,

the momentum for internationalizing its backbone, the ICANN, became greater. Recall that these pushes for internationalization are due in part to the perception of US government control over ICANN via the DoC and NTIA, shaped by the history of special relationships between state telecommunication corporations existing in other countries.

The (Potential) Tyranny of the ITU over Critical Internet Resources

One battleground for debates over internationalizing the ICANN was observed during preparations for the World Summit for the Information Society (WSIS),⁵⁴ when significant opposition to the current Internet governance began to emerge.⁵⁵ For instance, in March 2004 during a UN-hosted Global Forum on Internet Governance.⁵⁶ Brazilian delegate Maria Luiza Viotti claimed that Internet governance needed reform, since it is not inclusive of developing countries and instead appears to be under the ownership of one group of countries or stakeholders.⁵⁷ Lyndall Shope-Mafole, chair of South Africa's National Commission, spoke on similar lines, arguing that the legitimacy of the ICANN's processes, rather than its functioning, was of most concern for developing countries.⁵⁸ Thus, after rigorous talks, delegates concluded on the basis of concerns from the developing world that the ICANN required further reform. Throughout the WSIS process, and continuing in other forums discussing Internet governance and global cyber security, Brazil has continued to be a vocal proponent against the US position in the ICANN. In 2011, India joined South Africa and Brazil in proposing to "operationalize the Tunis mandate" by

bearing in mind the need for a transparent, democratic, and multilateral mechanism that enables all stakeholders to participate in their respective roles, to address the many cross-cutting international public policy issues that require attention and are not adequately addressed by current mechanisms and the need for enhanced cooperation to enable governments, on an equal footing, to carry out their roles and responsibilities in international public policy issues pertaining to the Internet, India proposes the establishment of a new institutional mechanism in the United Nations for global Internet related policies, to be called the United Nations Committee for Internet-Related Policies (CIRP).⁵⁹

The CIRP idea has gained momentum within the developing world as a counter to the current technical management of the Internet. Indeed, it echoes closely Chinese concerns voiced by the China Organizational

Name Administration Center (CONAC) that “the U.S. government has the sovereign power to control the Internet resources. We therefore suggest making the computer security plan available for comment by all multistakeholders, for maintaining the security of cyber space is not a mission only for the U.S. government, and it cannot be accomplished by any single nation.”⁶⁰

From Russia, then prime minister Vladimir Putin stated,

The International Telecommunication Union is one of the oldest international organisations; it's twice as old as the United Nations. Russia was one of its co-founders and intends to be an active member. We are thankful to you for the ideas that you have proposed for discussion. One of them is establishing international control over the Internet using the monitoring and supervisory capabilities of the International Telecommunication Union (ITU).⁶¹

Thus, the United States faces a significant challenge within the ITU from autocratic regimes leading the developing world to move control of critical Internet resources toward a multilateral body. The underlying danger is a shift away from an Internet whose defining characteristic is the free flow of information toward a model in which the political agendas of non-democracies attempt to exert control over the flow of information. Hence, the United States and like-minded nations must surge diplomatically to ensure the character of the Internet remains free from the political control of a multilateral institution.

This diplomatic struggle for control of the Internet has also been occurring within various other forums, like the UN Commission on Science and Technology for Development. Suggestions being made on the issue include:

Establishment of an ad hoc working group under the Commission on Science and Technology for Development with a view to the development of an institutional design and road map to enhance cooperation on Internet-related public policy issues with the support of the Secretary-General . . .

Creation of a more permanent committee on international public policy issues pertaining to the Internet within the United Nations system, possibly modeled on the Committee on Information, Communications and Computer Policy of the Organization for Economic Cooperation and Development . . .

And more concretely, global policy questions should be addressed by an entity with global representation, such as the United Nations, and regional questions by entities with regional representation, such as the Council of Europe . . . [and] the participation of relevant organizations in discussions on Internet governance at the quadrennial ITU Plenipotentiary Conference, and the public review process and Governmental Advisory Committee of ICANN.⁶²

With the upcoming World Conference on Telecommunications in December 2012, such statements indicate that these ideas will resurface as part of the ITU effort to revise International Telecommunications Regulations (ITR) to include governance of next-generation critical Internet resources within the ITU's mandate and assume a greater role in Internet governance.⁶³

Making Internet governance open to intergovernmental processes could put US national security at risk, given the potential for less-than-responsible state actors to take the current privatized laissez-faire approach to governing the Internet and have nation-states and their corporate entities take control of governing critical Internet resources. This would not ensure DoD equities are protected in an environment where critical decisions on underlying technical standards and Internet operation would be left to national governments that are competing with the United States.

Shadow “DNS” Rising

As described above, the critical Internet resources that allow for universally resolvable URLs and global Internet communications are possible due to the root system that is managed by the ICANN and protocols designed, developed, and debated within the IETF (among other organizations). Although this allows for a free and open Internet to function, the standards and protocols that the ICANN uses to maintain the domain name registries can be used by individuals, ad hoc networks, and nation-states to design and deploy an alternative DNS system that can either be independent of or “ride on top” of the Internet. A corporate LAN, such as “.company-name” for internal company use, is an example of the first. When a group wishes to ride over the global DNS root but incorporate its own pseudo top-level domain, core operators of the pseudo domains can use specific software resources to resolve domains that are globally accessible within their alternative DNS system. American audiences can experience what it is like to enter an alternative DNS universe via the Onion router (TOR) network. Downloading the Onion router package and navigating to websites one would prefer to visit anonymously (the typical use of TOR), one may point the TOR browser to websites on the “.onion” domain and mingle where the cyber underworld has started shifting the management of its business operations these days to avoid law enforcement and to add another layer of protection to their personas.⁶⁴

Should significant usage of such shadow Internets occur, this could lead to the loss of confidence and utility of the Internet itself. The greatest risk

comes when nation-states develop and deploy their own alternate domain-naming systems for internal use, thereby separating themselves from the global Internet. This is different from controlling access points and actually develops country-level intranets that may or may not be connected to the global Internet.⁶⁵ The discussion herein focuses on Russia and China as far as their successes in deploying potentially new intranets for in-country use. Other countries, such as Iran, are following suit.

US involvement in *openly* promoting and organizing “digital activists” by issuing up to \$30 million in grant funding to increase open access to the Internet, support digital activists, and push back against Internet repression wherever it occurs in the fight for free flows of information, generates international friction that is counterproductive to promoting international cooperation on cyber security issues.”⁶⁶ The “Internet Freedom Agenda” is one example of this phenomenon.⁶⁷ Such technology effectively allows citizen-activists to hack past government digital sentries to spread forbidden information. Other tools allow activists to don digital disguises and organize themselves into social movements designed to topple regimes. The result has been the emergence of alternative national networks that essentially create alternate domain name systems for in-country use, allowing for censorship of content and stifling the productivity of the current Internet topology. China is one country that has implemented this on a national scale, and Iran is closely following suit.⁶⁸ Others are sure to follow these attempts. The rise of a splintered Internet will certainly change the character of the current Internet, with negative consequences for freedom and prosperity worldwide. Those who wish the Internet to remain free and open will benefit, and draw a sharp, moral contrast with those wishing to control the master switch. Thus, maintaining the current Internet governance model, while addressing legitimate concerns of friends and allies, will help assure the Internet continues to serve as a robust platform for human economic development.

Conclusion

Failure to pay attention to our vulnerabilities from Internet governance and friendly conquest may provide our adversaries with a strategic advantage in cyber conflict. Our own cyber-attack efforts will also become complicated as networks that are not based on protocols and standards developed by US-based entities are deployed by our competitors. To aid

how we conceive of cyberspace, as well as adjust to change within the cyber environment, there must be a broad dialogue on these issues. Despite the Internet's historic roots within the Department of Defense, there has not been a well-organized effort to influence the development of technical standards and policies affecting Internet governance. Currently, the DoD has remained in a reactive mode, coordinating and commenting on the various global norms and standards being considered within the USG processes related to Internet governance. Because of this approach, the DoD and the USAF may be perceived as not having the legal expertise or technical reputation in Internet governance. The DoD, and the US Air Force in particular, should exercise leadership and take a more active role in the development of information technology infrastructure standards as it once did. Furthermore, it should more carefully document its role and provide metrics on its participation and position with Internet governance bodies. The Air Force should play a leading role within the DoD and the whole of government by explicitly focusing on a broader concept of friendly conquest that implicitly exists in policies, strategies, and doctrines. The 2012 World Telecommunications Conference in December 2012 may be the right place to commence this effort.

As the hardware and software on which the global Internet is based evolve and non-US entities begin to invent new hardware, standards, and protocols, potentially taking market share away from US entities, the US position as core cyber infrastructure operator will diminish. The United States currently enjoys technological dominance through its position of developer and core provider of Internet services made possible by the ICANN and the top-level Domain Name System. But our national cyber security strategies do not adequately address threats that may stem from other countries developing the protocols, standards, and technologies on which the next generation of networks will be based. The Air Force has a key role to play given the wealth of technical excellence that resides within its community of scientists and engineers. It cannot act alone, however, and the DoD will need to focus some of its already limited cyber resources toward Internet governance. Not doing so risks allowing foreign-designed technical standards and protocols to form the backbone of next-generation IT and potentially puts DoD operations at risk by reversing what is now an Internet characterized by the free flow of information on which the DoD depends. The USAF remains the leading

US military service impacting cyberspace, and thus its actions or inactions in Internet governance debates matter. **SSQ**

Notes

1. John Perry Barlow, "A Declaration of Independence of Cyberspace," published online 8 February 1996.

2. Tim Wu, *The Master Switch: The Rise and Fall of Information Empires* (New York: Alfred A. Knopf, 2010), 290.

3. Granted, for the most part, manufacturing does not occur within the United States, which presents the national security risk of supply-chain vulnerabilities. This is a subset of friendly conquest but remains beyond the scope of the argument here.

4. American values are a core national interest. *National Security Strategy* (Washington: The White House, May 2010), 35.

5. See, for example, Bryan Krekel, Patton Adam, and George Bakos, *Occupying the Information High Ground: Chinese Capabilities for Computer Network Operations and Cyber Espionage* (Washington: US-China Economic and Security Review Commission, 27 March 2012); and Dmitri Alperovitch, *Revealed: Operation Shady RAT* (Santa Clara, CA: McAfee White Paper, 2011), <http://www.mcafee.com/us/resources/white-papers/wp-operation-shady-rat.pdf>.

6. Martin Libicki, *Conquest in Cyberspace* (New York: Cambridge University Press, 2007).

7. The *National Strategy to Secure Cyberspace (NSSC)* (Washington: The White House, February 2003); John Rollins and Anna C. Henning, *Comprehensive National Cybersecurity Initiative (CNCI)* (Washington: Congressional Research Service, 10 March 2009; declassified in March 2010); the *International Strategy for Cyberspace* (Washington: The White House, May 2011); and the *Department of Defense Strategy for Operating in Cyberspace* (Washington: DoD, July 2011) are to date the leading relevant directives on cyber security. Although the White House completed a cyberspace policy review in 2009, the primary suggestions in the review amount to existing policy recommendations already in the *NSSC* and declassified *CNCI*. After the White House *Cyberspace Policy Review*, several initiatives were either launched or announced by departments and agencies of the US government. Declassification of the *CNCI* enabled the timely development of a framework for international partnerships consistent with a common cyber security policy. In 2011, the White House released the *International Strategy for Cyberspace*. Subtitled, *Prosperity, Security, and Openness in a Networked World*, the document falls short of providing the solutions necessary to live up to its name. The simple fact is, without security there can be no prosperity or openness.

8. *International Strategy for Cyberspace*, 12.

9. *Ibid.*, 15.

10. Lawrence Lessing, "Code is Law," in *Code: And Other Laws of Cyberspace, Version 2.0* (New York: Basic Books, 2006), 11–10.

11. Laura DeNardis, *Protocol Politics: The Globalization of Internet Governance* (Cambridge: MIT Press 2009), 11.

12. Cyrus Farivar, "Security Researcher Unearths Plans for Iran's Halal Internet," *Ars Technica*, 17 April 2012, <http://arstechnica.com/tech-policy/2012/04/iran-publishes-request-for-information-for-halal-internet-project/>.

13. *Department of Defense Information Enterprise Strategic Plan 2010–2012* (Washington: DoD, May 2010), 10, <http://dodcio.defense.gov/Portals/0/Documents/DodIESP-r16.pdf>. Examples con-

tained in the plan include the Internet Engineering Task Force, ICANN, Internet Governance Forum, Réseaux IP Européens, and American Registry for Internet Numbers/North American Network Operators' Group.

14. Libicki, *Conquest in Cyberspace*, 12.

15. *Ibid.*, 137.

16. The global positioning system (GPS) is one example where control of both the software and hardware is being contested. Although access to GPS is available without a fee for the basic service, friends and competitors alike have realized their dependence on this US system makes them vulnerable. Russia is modernizing its GPS system, and the European Union and China are developing independent GPS systems of their own. The long time cycle from intent to implementation of these new systems is due to the immense financial costs of deploying a space network. Cyber time cycles may be shorter, given the lower costs associated with deploying a national computer network compared with multiple high-tech satellites launched into space. For a more complete discussion of alternate GPS systems, see Lt Col Scott W. Beidleman, *GPS versus Galileo: Balancing for Position in Space* (Maxwell AFB, AL: Air University Press, 2006).

17. Air Force Doctrine Document (AFDD) 3-12, *Cyberspace Operations*, 2010, 4.

18. Bruce Rayner, "Ferretting out the Fakes," *Electronic Engineering Times*, 15 August 2011, 24. See also John Markoff, "Computer Gear may Pose Trojan Horse Threat to Pentagon," *New York Times*, 10 May 2008, 12.

19. *The National Security Implication of Investments and Products from the People's Republic of China in the Telecommunications Sector*, U.S.-China Economic and Security Review Commission Staff Report, January 2011, 7, http://www.uscc.gov/RFP/2011/FINALREPORT_TheNationalSecurityImplicationsofInvestmentsandProductsfromThePRCintheTelecommunicationsSector.pdf.

20. LCDR A. Anand, "Threats to India's Information Environment," in *Information Technology: The Future Warfare Weapon* (New Delhi: Ocean Books Pvt. Ltd., 2000), 56–62.

21. "Huawei Conducts World's First Commercial Network LTE Category 4 Trial," *Cellular News*, 9 May 2012, <http://www.cellular-news.com/story/54329.php>.

22. "China Information Technology Security Certification Center Source Code Review Lab Opened," *Microsoft News Center*, 26 September 2003, <http://www.microsoft.com/presspass/press/2003/sep03/09-26gspchpr.msp>.

23. Barrington M. Barrett Jr., "Information Warfare: China's Response to U.S. Technological Advantages," *International Journal of Intelligence and Counterintelligence* 18, no. 4 (2005): 682–706.

24. *Ibid.*

25. Clay Wilson, *Avatars, Virtual Reality Technology, and the U.S. Military: Emerging Policy Issues* (Washington: Congressional Research Service, April 2008), 4, 12.

26. Anton A. Huurdeman, *The Worldwide History of Telecommunications* (Hoboken, NJ: John Wiley & Sons, 2003), 91–146, 153–85. See also Jill Hills, "International Market Structure and the ITU," in *Telecommunications and Empire* (Champaign: University of Illinois Press, 2007), 91–116.

27. Edward Comor, "Communication Technology and International Capitalism: The Case of DBS and US Foreign Policy," in *The Global Political Economy of Communication: Hegemony, Telecommunication and the Information Economy*, ed. Comor (New York: St Martin's Press, 1994), 83–102.

28. Jill Hills, *The Struggle for Control of Global Communications: The Formative Century* (Champaign: University of Illinois Press, 2002.)

29. Parminder Jeet Singh, "India's Proposal Will Help Take the Web out of U.S. Control," *Hindu Online*, 17 May 2012, <http://www.thehindu.com/opinion/op-ed/article3426292.ece>.

30. Debora L. Spar, "National Policies and Domestic Politics," in *The Oxford Handbook of International Business*, ed. Alan M. Rugman (New York: Oxford University Press, 2008), 207.

31. *Ibid.*, 209.

32. *Ibid.*, 212.

33. Standard export restrictions are meant to prevent access, whereas sanctions or embargoes aim to act as punitive measures. Sanctions appear to have the greatest effects on firms. For example, firms in State I which imports from State A will be at a loss if State A subjects State I to a sanctions regime. However, firms that export from State A to State I will also be at a loss since they will suffer from a decline in sales and face the possibility of ties being severed with State I in the long term. Thus, as Spar notes, MNCs must remain aware of political developments within the countries in which they operate so as to not find themselves prohibited from accessing a market due to sanctions. Thus, export controls are one mechanism that can affect the behavior of firms and economies.

34. Richard C. Barth and Clint N. Smith, "International Regulation of Encryption: Technology Will Drive Policy," in *Borders in Cyberspace: Information Policy and the Global Information Infrastructure*, eds. Brian Kahin and Charles Nesson (Cambridge: MIT Press 1998), 283–99.

35. James Bamford, *The Shadow Factor: The Ultra-Secret NSA from 9/11 to the Eavesdropping on America* (New York: Doubleday, 2009). See also Claude Crépeau and Alain Slakmon, "Simple Backdoors for RSA Key Generation," in *CT-RSA'03: Proceedings of the 2003 RSA conference on the Cryptographers' Track* (Berlin: Springer-Verlag, 2003), 403–16; and Benjamin J. Romano, "Microsoft Device Helps Police Pluck Evidence from Cyberscene of Crime," *Seattle Times*, 29 April 2008, http://seattletimes.nwsourc.com/html/microsoft/2004379751_msftlaw29.html.

36. See Foreign Intelligence Surveillance Act, Electronic Communications and Privacy Act, and Communications Assistance for Law Enforcement Act.

37. The crux of the argument made by those holding the opinion that states' sovereignty is at bay is that "the multinational corporation has broken free from its home economy and has become a powerful independent force determining both international and political affairs. [While] others [who] reject this argue that the multinational corporation remains a creature of its home economy." It follows that by the MNC breaking free from its home economy, the sovereignty and autonomy of states is compromised. Those that disagree with the above claim argue that the MNC has not become fully independent from the home country but remains "a creature of the home country." Robert Gilpin, *Global Political Economy: Understanding the International Economic Order* (Princeton, NJ: Princeton University Press, 2001), 278.

38. Department of Commerce, *Management of Internet Names and Addresses*, 63 Fed. Reg. 31741 (1998).

39. Harold Kwalwasser, "Internet Governance," in *Cyber Power and National Security*, eds. Franklin D. Kramer, Stuart H. Starr, and Larry K. Wentz (Washington: NDU Press, 2009), 491–524.

40. DeNardis, *Protocol Politics*, 11.

41. See, for example, M. Ford et al., "Issues with IP Address Sharing," Internet Engineering Task Force, Request for Comments: 6269, June 2011, <http://www.hjp.at/doc/rfc/rfc6269.html>.

42. Ingrid Marson, "China launches largest IPv6 network," *CNET News*, 29 December 2004, http://news.cnet.com/China-launches-largest-IPv6-network/2100-1025_3-5506914.html.

43. Sheila Frankel et al., *Guidelines for the Secure Deployment of IPv6* (Gaithersburg, MD: National Institute of Standards, December 2010).

44. Katherine Kebisek, "AFNIC prepares Air Force for IPv6 transition" Air Force Space Command, 4 April 2011, <http://www.afspc.af.mil/news1/story.asp?id=123249968>.

45. Indeed, one should recall that the World Wide Web, the commercial adaptation of the DARPA-net project, was a CERN (European Organization for Nuclear Research) initiative.

46. This agreement was renewed on 2 July 2012. See <http://www.icann.org/en/news/announcements/announcement-2-09jul12-en.htm>.

47. Robert E. Molyneux, *The Internet under the Hood: An Introduction to Network Technologies for Information Professionals* (Westport, CT: Libraries Unlimited, 2003), 86.

48. ICANN, "Memorandum of Understanding Concerning the Technical Work of the Internet Assigned Numbers Authority," 1 March 2000, <http://www.icann.org/en/general/ietf-icann-mou-01mar00.htm>.

49. Elihu Zimet and Edward Skoudis, "A Graphical Introduction to the Structural Elements of Cyberspace," in *Cyber Power and National Security*, 91–112. See also Molyneux, *Internet under the Hood*, 85–86.

50. Molyneux, *Internet under the Hood*, 27.

51. The Internet governance debates have a history of about a decade and certainly will continue past 2012. The next phase of the World Summit for the Information Society will occur in 2015.

52. H. Zhao, "ITU and Internet Governance—Input to the 7th meeting of the ITU Council Working Group on WSIS, 12–14 December 2004," <http://www.itu.int/ITU-T/tsp-director/itut-wsis/files/zhao-netgov02.doc>.

53. For a comprehensive discussion of the dynamics of Internet politics as they relate to the perceptions by foreign countries that ICANN control is a cyber security for all, see Panayotis A. Yannakogeorgos, "Cyberspace: The New Frontier and the Same Old Multilateralism," in *Global Norms, American Sponsorship, and the Emerging Pattern of World Politics*, ed. Simon Reich (New York: Palgrave, 2010).

54. The World Summit on the Information Society (WSIS) and its spin-off, the Internet Governance Forum (IGF), are the main venues where governments and all interested stakeholders debate the issues, determine the objectives, and determine principles surrounding the structure of the global information society. The first and second phases of the summit resulted in the *Geneva Declaration of Principles* and the *Tunis Plan of Action*, respectively.

55. The ITU is the main entity tasked with organizing the WSIS. The High-Level Summit Organizing Committee was formed to "coordinate the efforts of the United Nations family in the preparation, organization and holding of WSIS." It was made up of a representative of the UN secretary-general and the executive heads of relevant UN specialized agencies. Other UN entities were included as observers. The ITU secretary-general served as the chair of this committee. One of its important functions was to "ensure that the contributions of the actors participating in the various conferences were comprehensively merged with the contributions from preparatory committees and regional meetings in a consensus document that would serve as the basis for the *Declaration of Principles* and *Plan of Action* of the WSIS."

56. "UN ICT Task Force Global Forum on Internet Governance to be Held in March," UN press release, Paris, 13 February 2004, http://portal.unesco.org/ci/en/ev.php-URL_ID=14347&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html.

57. "Global Internet Governance System is Working But Needs to Be More Inclusive, UN Forum on Internet Governance Told," UN press release, 26 March 2004, <http://www.un.org/News/Press/docs/2004/pi1568.doc.htm>.

58. Ibid.

59. "Statement by Mr. Dushyant Singh, Member of Parliament, on Agenda Item 16—Information and Communication Technologies for Development, at the 66th Session of the United Nations General

Assembly on October 26, 2011,” <http://content.ibnlive.in.com/article/21-May-2012documents/full-text-indias-un-proposal-to-control-the-internet-259971-53.html>.

60. Yang Yu, Chinese response to “Further Notice of Inquiry on the Internet Assigned Numbers Authority Functions,” China Organizational Name Administration Center (CONAC), http://www.ntia.doc.gov/files/ntia/conac_response_to_fnoi.pdf. CONAC is a nonprofit organization established in 2008. With the authorization of the State Commission Office for Public Sector Reform (SCPSR) and the Ministry of Industry and Information Technology (MIIT), CONAC runs the registry for “.政务.cn” (Government Affairs) and “.公益.cn” (Public Interest). CONAC also actively participates in the global Internet community.

61. “Prime Minister Vladimir Putin meets with Secretary General of the International Telecommunication Union Hamadoun Toure,” *Working Day*, 15 June 2011, <http://premier.gov.ru/eng/events/news/15601/>.

62. UN General Assembly, “Enhanced Cooperation on Public Policy Issues Pertaining to the Internet,” Report of the Secretary-General, http://unctad.org/meetings/en/SessionalDocuments/a66d77_en.pdf.

63. Signed by 178 countries, the ITR is a global treaty applied around the world.

64. Disclaimer: This is for informational use only. Any action undertaken by the reader of this article on the .onion domain is at his/her own risk, and this author is not liable for any harm caused by or to the reader.

65. This is different from what Chris Demchak points to in “Rise of a Cybered Westphalian Age,” *Strategic Studies Quarterly* 5, no. 1 (Spring 2011): 32–61, where the focus on sovereignty of the Internet is on access points of incoming Tier 1 ISP connections into the country and maintaining government control of those.

66. US Department of State, “Internet Freedom Fact Sheet,” 15 February 2011, <http://www.state.gov/r/pa/prs/ps/2011/02/156623.htm>.

67. Spencer Ackerman, “Does Obama’s ‘Net Freedom Agenda’ Hurt the U.S.?” *Wired*, 28 January 2011, <http://www.wired.com/dangerroom/2011/01/does-obamas-internet-freedom-agenda-hurt-the-u-s-without-helping-dissidents/>.

68. Ye Tian et al., “China’s Internet: Topology Mapping and Geolocating,” <http://cis.poly.edu/~ratan/topologymappingchinainternetshort.pdf>.

The Customary International Law of Cyberspace

Gary Brown, Colonel, USAF

Keira Poellet, Major, USAF

The first thing to know about international law is that it bears only a passing resemblance to the kind of law with which most people are familiar. Domestic laws in most countries are passed by some sort of sovereign body (like Congress) after due consideration. Statutes are carefully crafted so the law has a precise effect. International law is nothing like that. Contrary to popular belief, treaties are not the primary means of establishing international law. The body of international law is a jumble of historic practice and tradition as well as signed agreements between nations.

Within this patchwork of guidance, customary international law occupies a position of preeminence in developing areas of the law—ahead of treaties and conventions.¹ Customary international law develops from the general and consistent practice of states if the practice is followed out of a sense of legal obligation.² When this occurs, customary law is considered legally binding on nation-states. In situations not addressed by established consensus on what constitutes lawful behavior, nations may take actions they deem appropriate.³ This is the heart of the well-established *Lotus* principle, so named for the International Court of Justice decision in which it was established.⁴

Only a handful of actions are considered peremptory norms of international law; that is, things that are universally held to be wrong and impermissible.⁵ These are exceptional areas, including piracy, human trafficking, and hijacking. One reason there are so few universally accepted norms is the very nature of the international legal regime. It is established

Col Gary Brown has been the staff judge advocate (SJA) at US Cyber Command, Fort Meade, Maryland, since its establishment in 2010. Previously, he was the SJA at Joint Functional Component Command—Network Warfare. He is a graduate of the University of Nebraska College of Law.

Maj Keira Poellet is an operations law attorney at US Cyber Command. Her previous assignment was deputy SJA at Lajes Field, Azores, Portugal. She received her LLM in space and telecommunications law from the University of Nebraska College of Law and her JD from Whittier Law School.

by what nations do and believe they are bound to do, making consensus difficult to reach. Without consensus, there is no law, even in what seem to be straightforward cases, such as torture. “Torture or cruel, inhuman, or degrading treatment or punishment” is recognized by most states as violating human rights principles that have attained the status of customary international law. Yet, actions amounting to torture continue, and states sponsoring those actions are not often condemned, so it cannot be said there is complete international agreement on the issue.⁶

Although the few prohibitions accepted as peremptory norms do not deal with war, that is not to say armed conflict is completely ungoverned. There is a body of customary law reflecting the extensive and virtually uniform conduct of nation-states during traditional warfare that is widely accepted and well understood—the law of war. Unfortunately, the application of the law of war to cyberspace is problematic because the actions and effects available to nations and nonstate actors in cyberspace do not necessarily match up neatly with the principles governing armed conflict. Cyberspace gives nation-states new options, enabling them to take non-kinetic actions that may not have been available previously. Actions that may have required the use of military force in previous conflicts now can be done with cyber techniques without the use of force. States can also take actions in cyberspace that would be consistent with the use of armed force but more easily avoid taking responsibility for the actions—they can take cyber action “without attribution.”

In the absence of a specific legal regime for cyberspace, the logical approach is to take what guidance exists to govern more conventional warfare and determine whether it can be applied to cyberspace activities. The subsequent brief discussion is a general examination of how national practices become customs binding on the body of nations as customary international law. Following the general discussion is a more detailed discussion of how customary international law might apply to nation-state cyber actions.

The Development of Customary International Law

It is common for states to disagree about what constitutes a general practice accepted as law. The easiest form of proof is found in state actions, published government materials, official government statements, domestic

laws, and court decisions that detail actual practice.⁷ Over time, specific instances of state practice may develop into a general custom.⁸

The second part of the equation is more difficult. For a custom to be binding, states not only need to act in a certain way; they have to act that way because they think they are legally obligated to do so.⁹ Acceptance of general practice as an obligation, that it is “accepted by law,” is referred to as *opinio juris*.¹⁰ Evidence of *opinio juris* is primarily shown through statements of belief, as opposed to statements about state practice, such as treaties or declarations.¹¹

There is no mathematical formula governing how many states must accept a practice or for how long it needs to be practiced for it to become binding custom.¹² For the most part, the more states that practice a custom, the more likely it is to evolve into law, but not even that simple rule holds completely true. The practice of politically powerful and active states carries more weight than that of smaller nations, especially ones not actively engaged in the area under consideration. For example, actions of the United States or Great Britain will have more bearing on the development of international law governing naval operations than those of Switzerland.

As noted, the length of time to develop customary international law can vary greatly. The law of war is a good example. The customary law of war has developed over thousands of years, but the practice of limiting conflict (e.g., to protect noncombatants) evolved primarily in the last 150 years. For example, the Greeks began developing the concept of *jus ad bellum*, or just war, in the fourth century BC.¹³ By contrast, while the principles governing the way in which combatants engage in warfare (*jus in bello*) also have historical ties to that era, they did not begin to assume their current form until the 1860s during the Franco-Prussian War and the American Civil War. Documented atrocities during those wars led to rapid development of the modern law of war regime, beginning with the first Hague Convention in 1899.

An example of customary law that developed quickly is space law.¹⁴ In 1958, just one year after the launch of *Sputnik*, the UN General Assembly created a committee to settle on the peaceful uses of outer space. By 1963, the United Nations had put forth the *Declaration of Legal Principles Governing the Activities of States in the Exploration and Use of Outer Space*, formally recognizing what had become customary law applicable to space activities. Since then, most space law has been generated through international agreements, beginning with the first outer space treaty signed in 1967.

Sometimes even state *inaction* can establish practice. For example, when one state engages in conduct harmful to another, the official silence of the “victim” state can be evidence that the conduct in question does not constitute a violation of international law. This passiveness and inaction can produce a binding effect under what is called the doctrine of acquiescence.¹⁵ The more times a state permits an action to occur without meaningful protest, the more likely it is the action will be accepted as lawful state practice.

Development of Cyber Law through Custom

The increasing use of computers and computer networks through the 1970s and 1980s was followed swiftly by the rise of the “network of networks” known as the Internet in the mid-1990s.¹⁶ Ultimately, the Internet spawned an entirely new domain of operations referred to as *cyberspace*. It is in and through this virtual space that cyber activities occur. So, not only are the activities in cyber new, *where* cyber actions take place is a unique location.¹⁷

Because it has existed for such a short time, there is not a robust body of law governing state conduct in cyberspace.¹⁸ There are documented instances of state cyber practice, however, and these have begun to lay a pattern for establishing customary cyber law. As noted above, customary law does not instantly appear but is developed through state practice and rationale. The cyber practices of states and the thought behind those actions over the past 30 years must be examined to determine if there is customary law in cyberspace. If no principles have developed, as earlier discussed, cyberspace remains unconstrained under the default customary international regime.

Although *opinio juris* is a critical element, it is easiest to analyze the development of custom beginning with an examination of state action, which is more visible and easily documented than motivation. Complicating the analysis is the secrecy surrounding most cyber operations. The US Department of Defense (DoD), for example, claims it suffers millions of scans and thousands of probes into its networks each day.¹⁹ With rare exceptions, no states or individuals come forward to take credit for these actions, so assessing the motivation of these unknown cyber actors is difficult. Albeit complicated and difficult, a few examples of state practice in cyber are available for examination.

Arguably, the first cyber attack occurred in the Soviet Union. In 1982, a trans-Siberian pipeline exploded. The explosion was recorded by US satellites, and it was referred to by one US official as “the most monumental nonnuclear explosion and fire ever seen from space.”²⁰ It has been reported the explosion was caused by computer malware the Central Intelligence Agency implanted in Canadian software, apparently knowing the software would be illegally acquired by Soviet agents. Because the explosion happened in remote Siberia, it resulted in no casualties. It also embarrassed the Russian Committee for State Security (the KGB), who thought they had stolen the most recent software technology from the United States. As a result, the facts behind the explosion were concealed, and the USSR never publicly accused the United States of causing the incident.²¹

Multiple “soft” computer attacks occurred against US systems as the Internet grew exponentially over the next 25 years. Many of these involved attempts to copy sensitive information or relatively simple but potentially devastating denial of service attacks.²² Some of the more infamous include Moonlight Maze (1998–2001), which probed government and academic computer systems in the United States; Code Red (2001), which launched a worm intended to conduct a denial of service attack against White House computers; and Mountain View (2001), a number of intrusions into US municipal computer systems to collect information on utilities, government offices, and emergency systems.²³ Although there was speculation about the origins, none of these incidents could be definitively attributed to a state actor.

In contrast to the, until recently, little-known Siberian incident, it was a very public series of cyber events considered by many to have heralded the advent of cyber warfare. In April 2007, following the removal of a Russian statue in Estonia’s capital of Tallinn, a widespread denial of service attack affected its websites. As a result Estonia, one of the world’s most wired countries, was forced to cut off international Internet access. Russia denied involvement in the incident, but experts speculate the Russian Federal Security Service (FSB) was behind the distributed denial of service event.²⁴

The following year, Russian troops invaded the Republic of Georgia during a dispute over territory in South Ossetia. In August 2008, prior to Russian forces crossing the border, Georgian government websites were subjected to denial of service attacks and defacement. While there is widespread belief the incident was “coordinated and instructed” by elements

of the Russian government, no one has been able to attribute these actions definitively to Russia.²⁵

The wakeup call for the US military occurred in 2008, although the details did not become public until two years later. Operation Buckshot Yankee was the DoD's response to a computer worm known as "agent.btz" infiltrating the US military's classified computer networks.²⁶ The worm was placed on a flash drive by a foreign intelligence agency, from where it ultimately made its way to a classified network. The purpose of the malware was to transfer sensitive US defense information to foreign computer servers.²⁷ In what qualifies as bureaucratic lightning speed, US Cyber Command was established less than two years later, with a mission to, among other things, direct the operations and defense of DoD computer networks.²⁸ In addition to unmasking the extent of network vulnerabilities, the event highlighted the lack of clarity in international law as it relates to cyber events.

Two recent incidents merit attention before discussing the law in depth. In 2010, Google reported Chinese hackers had infiltrated its systems and stolen intellectual property. Through its investigation, Google learned the exfiltration of its information was not the only nefarious activity; at least 20 other companies had been targeted by Chinese hackers as well. These companies covered a wide range of Google users, including the computer, finance, media, and chemical sectors. The Chinese had also attempted to hack into G-mail accounts of human rights activists and were successful in accessing some accounts through malware and phishing scams. Google released a statement explaining what it discovered through its investigation and what steps it was taking in response to China's action, including limiting its business in and with China.²⁹

Also in 2010, a computer worm named Stuxnet was detected on computer systems worldwide. Stuxnet resided on and replicated from computers using Microsoft's Windows operating system but targeted a supervisory control and data acquisition (SCADA) system manufactured by Siemens. Cyber experts determined the worm was designed to affect the automated processes of industrial control systems and speculated that either Iran's Bushehr nuclear power plant or its uranium enrichment facility at Natanz was the intended target.³⁰ After Stuxnet became public, Iran issued a statement that the delay in the Bushehr plant becoming operational was based on "technical reasons" but did not indicate it was because of Stuxnet.³¹ The deputy director of the Atomic Energy Organization of Iran stated,

“Most of the claims made by [foreign] media outlets about Stuxnet are efforts meant to cause concern among Iranians and people of the region and delay the launch of the Bushehr nuclear power plant.”³² Iranian president Ahmadinejad stated at a news conference that malicious software code damaged the centrifuge facilities, although he did not specifically state it was Stuxnet or the Natanz facility.³³

Even disregarding the Siberian pipeline incident and considering Moonlight Maze the first major state-on-state cyber incident, there have been about 12 years of general practice to consider when determining what constitutes customary law in cyberspace. Incidents that have occurred during this period have set precedent for what states consider acceptable cyber behavior. What is remarkable is the lack of protest from nations whose systems have been degraded in some way by obnoxious cyber activity. Iran seemed reluctant even to admit its nuclear plant’s computers had been affected and still does not claim to have been cyber attacked.³⁴

If the damage caused by the Stuxnet malware had instead been caused by a traditional kinetic attack, such as a cruise missile, it is likely Iran would have vigorously responded. For one thing, in more-traditional attacks it is easier to determine the origin of attack. There are a variety of reasons Iran may have refrained from public complaint over the Stuxnet event; one possibility is that it believes the action was not prohibited under international law. Whatever the reason for Iran’s silence, it remains true that no state has declared another to have violated international law by a cyber use of force or an armed attack through cyberspace. Aside from the Stuxnet event, those in Estonia and Georgia came closest.

The situation in Georgia can be distinguished because the cyber action was taken in concert with Russian troops crossing the Georgian border—a clear use of force. Cyber activity against Georgian websites did not start until after Georgia made its surprise attack on the separatist movement in South Ossetia on 7 August 2008. The cyber activity commenced later that same day, on the eve of Russia launching airplanes to bomb inside Georgian territory. It appears as though it was a military tactic to sever Georgia’s ability to communicate during the attack. It was not until 9 August 2008 that Georgia declared a “state of war” for the armed attack occurring inside its territory. It did not declare the cyber activity itself an attack or use of force.³⁵

A case has also been made that the 2007 massive distributed denial of service activity in Estonia was a cyber attack. However, after deliberation,

even the Estonian government concluded it was a criminal act as opposed to a use of force by another state. That may be because they were not able to attribute it with certainty to the Russian government (or any other government), but the precedent remains. Attribution problems will continue to plague this area of law. It is more difficult for custom to develop if the source of the action is unknown. The actions of criminal gangs or recreational hackers do not set precedent for international law, and as long as the actor remains unknown, the events have no precedential value.

Cyber Activity and Espionage

Much of what has occurred in cyberspace between states can be viewed as merely espionage—simply intrusions onto computer systems for the collection of intelligence. If these actions are equivalent to espionage, however, this creates a dilemma in the analysis of cyber law.

Spying has been around even longer than customary international law. Despite the famous statement, “Gentlemen do not read other gentlemen’s mail,” espionage has existed since the earliest days of armed conflict.³⁶ Although the law of war addresses wartime espionage and the treatment of captured spies, customary international law is notably silent on the practice of spying during peacetime. States have domestic laws prohibiting espionage—including the United States, where spying is punishable by death—but there is no international law prohibiting espionage or insisting it violates sovereignty.³⁷

Despite the absence of specific guidance, it is generally not argued that espionage is actually legal under international law. Most international lawyers contend espionage is “not illegal” internationally. Presumably, this is because it would be unseemly for countries to openly note that it is acceptable to undertake as much espionage as they can get away with. Despite the “ungentlemanly” nature of espionage, it is an open secret that countries spy on friends and foes alike. Most of the time, when spies are caught, the result is a declaration of “PNG” (*persona non grata*) and deportation or an exchange for other spies.³⁸

The practice of nations with regard to espionage amounts to a tacit acceptance of spying. The activity is not overtly endorsed but rather occupies an ill-defined policy space that permits it to occur without violating international law. There is a general prohibition against violating territorial sovereignty, but as an exception to the rule, state practice does not

prohibit spying that might involve crossing international borders without permission. Reflecting this general view, one author summarized, “The law of espionage is, therefore, unique in that it consists of a norm (territorial integrity), the violation of which may be punished by offended states, but states have persistently violated the norm, accepting the risk of sanctions if discovered.”³⁹

This assertion aptly illustrates the bizarre position espionage holds in the international community. Years of state practice accepting violations of territorial sovereignty for the purpose of espionage have apparently led to the establishment of an exception to traditional rules of sovereignty—a new norm seems to have been created. As cyber activities are frequently akin to espionage, even if conducted for another purpose, perhaps it is not too much of a leap to assert that most cyber activities can also occur without violating territorial sovereignty.

As states have begun to use the Internet and other computer capabilities to store, process, and communicate information, the use of cyber capabilities by intelligence agencies around the world has similarly increased. “Motives for spying [have not] changed in decades. What has changed are the means by which people spy. Cyber spying has accelerated due to increased network speeds and sophisticated chip processing capabilities.”⁴⁰ One might think this would mean all nonkinetic national cyberspace operations would be governed by the loose international standards of espionage. Unfortunately, it is not quite so simple.

Manipulating cyberspace in the interest of national security began with espionage, but the continuing development of cyber capabilities means it could be used in military operations independent from espionage. Perhaps for this reason, policies and practices governing cyber espionage are more fully developed than those governing official cyber activities undertaken for other reasons. Objectively, there is little rationale for this disconnect, as most military actions in cyber would fall short of a use of force. In fact, many military actions in cyber would be indistinguishable from cyber espionage.

On the other hand, in some cases there are important differences between cyber espionage and more traditional means of spying. Surreptitiously entering a foreign country and leaving behind a sensor to collect and transmit intelligence data is one thing. But what if that sensor also contained a powerful explosive that could be detonated from a distance, causing grave destruction? If a government discovered such a device, it would be classified as a weapon of war; that would subsume any thought that it might

have been placed during an espionage activity. This second scenario is perhaps more akin to some current cyber espionage techniques. Network accesses and cyber spying capabilities may be just as capable of being used for disruption of systems or deletion of data. The cyber victim may be left to wonder whether the rogue code it discovers on its network is a tool meant for espionage or attack.

A nation on the receiving end of espionage-like cyber activity (such as illicitly gaining access to a government computer network) has no sure method of discerning the intent of an intrusion and may have little notion of who is behind it. Whatever unauthorized access is gained through nefarious means could be used to collect data, destroy data, or even damage or destroy equipment. "The difference between cybercrime, cyber-espionage and cyberwar is a couple of keystrokes. The same technique that gets you in to steal money, patented blueprint information, or chemical formulas is the same technique that a nation-state would use to get in and destroy things."⁴¹ Once illegitimate users have access to a network, they can conduct whatever mischief they like, and the software tools used by spies might well be the same as those used by criminals and saboteurs.

So, even if the target government could effectively attribute the activity to a certain state, it would not know the "why" of the activity. The nature of cyberspace does not allow for a clear distinction between intrusions for collection means and those of a more nefarious nature.

For this reason, it might follow that cyberspace operations that fall below the use of force should be covered by the same broad international law umbrella of "not illegal" that governs espionage. After all, most military cyber activities are more similar to espionage than they are to traditional military action.⁴² Conceptually, there is little difference between tip-toeing into an office and stealing a sheaf of papers from a file cabinet and electronically sneaking into a computer to steal a file. There is a significant difference, however, between destroying something and a reversible action temporarily rendering something less functional. In the kinetic realm, few minimally invasive options are available. In cyber, options range from tweaking a single digit to crashing a national power grid. To treat all cyber activity equally as "attacks" is unreasonable.

To facilitate the collection of intelligence, computer code (malware) is planted in government systems. That code, in some cases, can either be used in intelligence gathering or in destructive ways, for example, to hard-break a computer system controlling e-mail at a military headquarters.

The system access created for intelligence purposes may also be used to disrupt computer systems at a level well below what would be considered a use of force under international law. Although it might be argued that the intent of the actor controls how a cyber action should be analyzed under international law, this line of argument tends to mix international and national standards of behavior.⁴³ A person's intent is key to many criminal charges under national law, yet in the law of war, a nation that feels threatened or as though it is under attack may not be especially concerned with the intent of the offending nation.

There is no international legal body to which states can turn to collect evidence and carefully analyze it to determine the intent behind another state's cyber activity. Neither the International Court of Justice nor other international courts can fill this role. Any evidence that existed would be classified as secret by the actor nation and would be politically sensitive as well. Witnesses would mostly be intelligence officials and politicians. In short, the system bears little resemblance to a national court system, where police officers, official reports, and witnesses may be scrutinized fully over the course of many months to determine intent. When a state becomes aware of a cyber intrusion, it must decide quickly whether it is a prelude to an attack or "merely" espionage. Even if the victim state were of a mind to inquire about intent, it might not be able to determine the source of the intrusion. Further, it might not want to disclose that it detected the intrusion.

The issue of international intent has not been much discussed as it applies under the law of war. That may be because, in the case of kinetic attacks, the intent of the attacking state is generally unambiguous.⁴⁴ This sets up an interesting conundrum. If intent does not matter in cyber operations, and only a few keystrokes determine whether a cyber activity will constitute espionage or attack, then any intrusion for collection purposes is potentially a threat or use of force. If that is the case, the UN Security Council could be set for a big increase in business.⁴⁵

The international legal system operates under its own rules, which are established by consensus and are fundamentally different than domestic law. The law of war is driven almost entirely by the effect of actions rather than by some sort of "national mens rea."⁴⁶ The *intent* of an actor taking an action against another state that could be interpreted as hostile is, for practical purposes, irrelevant to the international law analysis.

All this leads back to the current international legal regime governing cyber activities. The question is whether state practice coincides with these norms and whether states are complying out of a sense of legal obligation. Otherwise, it is still the “Wild West” when it comes to behavior in cyberspace.

In general, cyberspace is a permissive regime, analogous to the espionage rule set—little is prohibited, but states can still do their best to prevent others from playing in the arena. There is also nothing to prevent states from prohibiting cyber behavior with national laws. Specifically, as long as cyber activity remains below the level of a use of force and does not otherwise interfere with the target nation’s sovereignty, it would not be prohibited by international law, regardless of the actor’s intent.

One important caveat is that aggressive cyber activities resulting in kinetic effects (i.e., physical destruction, damage, or injury) are covered by the law regarding the use of force and armed attack. They are kinetic events, governed by the traditional law of war just like kinetic effects caused by more traditional means of warfare. So, for example, a cyber event resulting in the physical destruction of a power plant turbine would be a military attack subject to the same international law governing any other kinetic attack.⁴⁷ Although determining exactly what constitutes a kinetic effect is not always simple, this line is as clear as others governing the murky corners of customary law and is clear enough effectively to distinguish cyber attacks from something less. One example of the gray area is a cyber action against an electric power grid that causes it to temporarily cease functioning. Although no actual kinetic event may occur, the reliance of modern societies on electricity for health care, communications, and the delivery of essential services makes it clear this would qualify as a kinetic-like effect and would therefore constitute a military attack if the disruption were for a significant period of time.⁴⁸

Turning to areas of cyber operations that do not rise to the level of a military attack, there are few rules. But *few* is different than *none*, and some markers appear to have been set on the table to guide international attorneys in assessing the state of affairs.

In 2003, during the months leading up to the invasion of Iraq, the United States planned a cyber operation that would have greatly affected Iraq’s financial system and frozen billions of dollars during the opening stages of the war.⁴⁹ Ultimately, US officials chose to forego this option. Reportedly, this was because they were concerned an attack on one nation’s

financial system would affect international confidence in the global financial system, harming the United States and its allies as well as Iraq. So, there is some question about whether they refrained due to *opinio juris* or out of mere self-interest.

In the end, it makes little difference. The financial systems of modern states are inextricably intertwined, more now than in 2003. If any nation's action would most likely damage the financial systems of many other nations, it seems this type of action would be a violation of customary international law. If for no other reason, these actions would be questionable, as they would be indiscriminate. Financial systems include banking and stock markets, essentially any "high finance" connected to the international financial system. The worldwide recession of 2007–08 demonstrated again how when one of the world's large economies sneezes, the rest are likely to catch cold.⁵⁰

There is some potential counterevidence to this conclusion. In 2011, the NASDAQ reported an intrusion into its computer systems.⁵¹ NASDAQ is an important financial entity, and if shut down, would certainly qualify under our definition as a cyber attack; that is, a cyber activity that is impermissible under international law. In this case, however, it appears the intrusion was detected before any harm was done, and the United States may have decided it was criminal activity not meriting a diplomatic brouhaha, or NASDAQ may have been unable to determine the source of the penetration. This does not affect the conclusion here: large-scale disruption, or destruction, of a nation's financial institutions qualifies as cyber attack.

It also appears penetration or disruption of nuclear command and control systems is a violation of customary international law. This assertion is supported by the absence of state practice to the contrary and the abundance of *opinio juris* regarding the nonproliferation and the monitoring and control of nuclear weapons.⁵²

Other than these two areas, state cyber activity that falls below the level of a use of force is not prohibited under international law. It may be undertaken, just as espionage is, without sanction from the international community. Some examples of permissible behavior, as demonstrated by state practice, are penetrating and maintaining a cyber presence on government computer systems (including SCADA systems), exfiltration of government data (including the most sensitive military secrets), and denial of service or similar activities that decrease bandwidth available for government websites.

The above is premised on the thought that countries would react if they were attacked. Because all of these things have occurred but not elicited significant recriminations or a self-defense response, the conclusion is they are not attacks. However, those who take these actions in government systems run the risk of misperception that their cyber espionage is a cyber attack. If they are not armed attacks or uses of force under international law, they are not governed by the customary law of war. As a result, these disruptive cyber activities are governed by the overall customary law regime. As earlier discussed, the customary regime is permissive in the absence of norms, as is the case here. The closest existing analogy is to the rule set governing espionage. Under either the permissive or the espionage regime, disruptive cyber activities undertaken by states are permissible as a matter of customary international law, with the two exceptions (financial systems and nuclear command and control systems) noted here.

Shaping US Strategy for International Cyber Law

Because of its reliance on cyberspace, the United States should consciously craft a strategy to influence the development of customary international cyber law rather than merely observing the development. The best method to do so is through acknowledged state practice. Because of the secrecy involved in many cyberspace activities, few actually influence the development of norms. A prudent examination of US actions—and public disclosure of some—would help establish a baseline for acceptable behavior.

After the United States determines what actions it believes it is authorized to take in cyberspace, it should openly share at least examples of actions it has taken. Further, it should certainly look to the possibility of disclosing actions taken against it. By proposing certain of its own actions as acceptable and recognizing those taken against it as either acceptable or unacceptable, the United States could lead a dialogue on cyber norms, driving toward conclusions that would be beneficial for its national security.

In addition to state practice, the United States should provide releasable government materials stating what it believes are cyber norms. In May 2011 the president released the *International Strategy for Cyberspace*. This strategy recognizes that “the development of norms for state conduct in cyberspace does not require a reinvention of customary international law,

nor does it render existing international norms obsolete. Long-standing international norms guiding state behavior—in times of peace and conflict—also apply in cyberspace.”⁵³

In recognizing that certain principles apply to cyberspace activities just as they apply to more traditional activities, the United States provides a basic framework for the cyber norms it expects will develop: upholding fundamental freedoms, respect for property, valuing privacy, protection from crime, and the right of self-defense. Although at this point, the list is more aspirational than actual, it can serve as a framework on which the United States can hang future examples of real cyber behavior by itself and others.

It is important to note that the norms set out in the *International Strategy for Cyberspace* are not universally recognized as customary international law (except for the right of self-defense). For example, although the strategy discusses fundamental freedoms such as free speech and privacy, it is apparent that particular norm is not followed worldwide. Twitter, which has been an important communications tool for government protestors in many countries, announced that it will restrict certain speech and freedom of expression if it appears to violate a local law by “reactively withhold[ing] content from users in a specific country while keeping it available to the rest of the world.”⁵⁴ So, even if the United States does not, Twitter recognizes that not all these things are accepted as norms of behavior worldwide at this point.

The *Department of Defense Strategy for Operating in Cyberspace* (DSOC) recognizes the same principles and encourages the development and promotion of international cyberspace norms. The DSOC reiterates the *International Strategy*’s defense objective to “oppose those who would seek to disrupt networks and systems, dissuading and deterring malicious actors, and reserving the right to defend these vital national assets as necessary and appropriate.”⁵⁵ Neither strategy document includes actual examples of what would be necessary and appropriate and leaves it open to interpretation. While it is helpful to provide the statement that the United States has the right to defend its vital national assets, for the purpose of customary international law it would also be helpful to know what the United States considers as a threat to those assets. On the other hand, the United States may have intentionally left this ambiguity in its international strategy to allow for the flexibility of a relevant response.

Conclusion

In the absence of formal international agreements, cyber custom is beginning to develop through the practice of states. The custom permits most cyber activity that falls below the level of a use of force, with serious actions against major financial institutions and disruptive actions to nuclear command and control systems being notable exceptions. While there has been some movement toward declarations, agreements, treaties, and international norms in the area, the hopeful statements most often heard do not coincide with current state practice. In a practical demonstration of *realpolitik*, states generally would like to prohibit others from undertaking the same cyber activity in which they are already engaging. The disconnect between practice and public statements creates a poor environment for negotiating international agreements and infertile soil for positive customary law—norms—to flourish. In this case, for better or worse, the default—permissive international law regime—governs. Unless states positively determine that disruptive cyber actions should be treated differently than espionage, this area will continue to be a competitive intellectual battlefield, where the cyber savvy do what they will and the cyber naïve suffer what they must.

This is not necessarily a bad-news story. Recognizing the permissive nature of cyber custom will encourage states to negotiate agreements that moderate behavior in cyberspace. To negotiate agreements, states will have to address critical cyber issues of attribution and state responsibility. In the long run, negotiated and enforceable agreements governing cyberspace may be a better option than waiting for the necessarily languid development of custom in an area that changes at the speed of thought. **SSQ**

Notes

1. See *Statute of the International Court of Justice*, Art. 38 (18 April 1946), <http://www.icj-cij.org/documents/index.php?p1=4&p2=2&p3=0>.

2. John B. Bellinger III and William J. Haynes II, "A US Government Response to the International Committee of the Red Cross Study *Customary International Humanitarian Law*," *International Review of the Red Cross* 89, no. 866 (June 2007): 443–71, http://www.icrc.org/eng/assets/files/other/irrc_866_bellinger.pdf.

3. Guidance to the contrary may be exhibited, for example, through bilateral treaties or consistent objection by other states.

4. It is a "residual negative principle which provides that in the [absence of law], whatever is not prohibited in international law is permitted." Anthea Roberts, "Traditional and Modern Approaches to Customary International Law: A Reconciliation," *American Journal of Inter-*

national Law 95 (2001): 757–91. While it is possible that the *Lotus* principle could prompt states to attempt to regulate on any matter that could affect them negatively, international law expects that states “may not exercise jurisdiction to prescribe law with respect to a person or activity having connections with another state when the exercise of such jurisdiction is unreasonable.” *Restatement of the Law, Third, Foreign Relations Law of the United States*, §403, 1987 [hereinafter *Restatement*].

5. “A norm accepted and recognized by the international community of States as a whole as a norm from which no derogation is permitted and which can be modified only by a subsequent norm of general international law having the same character.” *Vienna Convention on Treaties*, Art. 53, 23 May 1969, http://untreaty.un.org/ilc/texts/instruments/english/conventions/1_1_1969.pdf.

6. The United States considers the prohibition on torture to be *jus cogens*, but as noted, the practice of nations may not support that conclusion. *Restatement*, §702, comment n.

7. *Restatement*, §102, comment b.

8. Roberts, “Traditional and Modern Approaches, 757–58.

9. *Restatement*, §102, comment c, n. 4. This comment also suggests that explicit evidence may not always be necessary to establish *opinio juris*; in some cases it may be inferred from state practice alone.

10. Peter Malanczuk, *Akehurst’s Modern Introduction to International Law*, 7th rev. ed. (London: Routledge, 1997), 39.

11. Roberts, *Traditional and Modern Approaches*, 758, n. 4.

12. *Restatement*, §102, comment b.

13. See Polybius, *The Histories*, Book V, 9 (discussing the right to retaliate for sacrilegious acts committed by Aetolians), http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Polybius/5*.html.

14. “The analysis of the practice of states before the conclusion of the 1967 Outer Space Treaty shows that historically, custom was the first source of the international law of outer space.” Vladelen S. Vereshchetin and Gennady M. Danilenko, “Custom as a Source of International Law of Outer Space,” *Journal of Space Law* 13, no. 1 (1985): 22, 25.

15. Malanczuk, *Akehurst’s Modern Introduction to International Law*, 43, n. 10. See I. C. MacGibbon, “The Scope of Acquiescence in International Law,” 1954 *British Yearbook of International Law*, 143, 145–46; and MacGibbon, “Customary International Law and Acquiescence,” 1957 *British Yearbook of International Law*, 115, 138.

16. Harry Newton, *Newton’s Telecom Dictionary*, 23rd ed. (New York: Flatiron Publishing, 2007), 502–3.

17. The DoD defines *cyberspace* as a war-fighting domain. Joint Publication 1-02, *DoD Dictionary of Military and Associated Terms*, 12 April 2001 (as amended through April 2010), 121.

18. As distinguished from state actions that use cyber capabilities merely as a means to accomplish a more traditional effect. For example, using e-mail to deliver a diplomatic note is legally no different than sending the note with the ambassador. The importance of “effects” is discussed below.

19. Deputy Secretary of Defense William J. Lynn III, “Remarks on Cyber,” Council on Foreign Relations, 30 September 2010, <http://www.defense.gov/speeches/speech.aspx?speechid=1509>.

20. Bret Stephens, “Long before There Was the Stuxnet Computer Worm, There Was the ‘Farewell’ Spy Dossier,” *Asian Wall Street Journal*, 19 January 2010, 10. In the early 1980s, a KGB officer leaked to French intelligence the names of Soviet agents involved in industrial espionage. This information was used by the West to feed misleading information to the USSR; the leaked data was referred to as the Farewell Dossier.

21. William Safire, "The Farewell Dossier," *New York Times*, 2 February 2004, <http://www.nytimes.com/2004/02/02/opinion/the-farewell-dossier.html?ref=williamsafire>.
22. A denial of service (DoS) attack prevents a website from being responsive by overwhelming it with thousands of requests (pings). Often these requests originate from a robotic network, more commonly referred to as a botnet. "Bots" are malware-infected computers belonging to unwitting individuals. The bots become part of a botnet—a grouping of bots—which is controlled by the unfriendly actor. Bots may be used to perform a variety of unsavory acts, such as sending spam and collecting data for identity theft. Botnets are usually composed of computers from many geographic locations, so the action is called a distributed DoS, or DDoS. Newton, *Newton's Telecom Dictionary*, 300, n. 16.
23. A worm is a type of computer virus that can spread without human action and duplicate itself through an entire network. A worm can allow an unauthorized user to remotely access a computer.
24. William Ashmore, "Impact of Alleged Russia Cyber Attacks," *Baltic Security and Defence Review* 11, no. 8 (2009).
25. Cooperative Cyber Defence Centre of Excellence (CCDCOE), *Cyber Attacks Against Georgia: Legal Lessons Identified* (Tallinn, Estonia: CCDCOE, November 2008), 12.
26. Noah Shachtman, "Insiders Doubt 2008 Pentagon Hack Was Foreign Spy Attack," *Wired: Danger Room*, 25 August 2010, <http://www.wired.com/dangerroom/tag/operation-buckshot-yankee/>; and Sergi Shevchenko, "Agent.btz: A Threat That Hit Pentagon," *Threat Expert* blog, 30 November 2008, <http://blog.threatexpert.com/2008/11/agentbtz-threat-that-hit-pentagon.html>.
27. William J. Lynn III and Nicholas Thompson, "Defending a New Domain," *Foreign Affairs* 89, no. 5 (September/October 2010).
28. US Cyber Command, "Mission Statement," <http://www.stratcom.mil>.
29. See Google's statement at <http://googleblog.blogspot.com/2010/01/new-approach-to-china.html>. Google has now resumed doing business in China.
30. Yossi Melman, "Computer Virus in Iran Actually Targeted Larger Nuclear Facility," *Haaretz.com*, 28 September 2010, <http://www.haaretz.com/print-edition/news/computer-virus-in-iran-actually-targeted-larger-nuclear-facility-1.316052>.
31. Ministry of Foreign Affairs, Islamic Republic of Iran, weekly briefing, 5 October 2010, <http://www.mfa.gov.ir/cms/cms/Tehran/en/NEW/137891.html>.
32. "No Delay in Launch of Bushehr Power Plant Due to Stuxnet: Official," *Tehran Times*, 5 February 2011, http://www.tehrantimes.com/index_View.asp?code=23518.
33. Mark Clayton, "Stuxnet: Ahmadinejad Admits Cyberweapon hit Iran Nuclear Program," *Christian Science Monitor*, 30 November 2010, <http://www.csmonitor.com/USA/2010/1130/Stuxnet-Ahmadinejad-admits-cyberweapon-hit-Iran-nuclear-program>.
34. See, for example, Bob Sullivan, "Could Cyber Skirmish Lead U.S. to War?" *Red Tape Chronicles*, 11 June 2010, <http://redtape.msnbc.com/2010/06/imagine-this-scenario-estonia-a-nato-member-is-cut-off-from-the-internet-by-cyber-attackers-who-besiege-the-countrys-bandw.html>; and Gary D. Brown, "Why Iran Didn't Admit Stuxnet Was an Attack," *Joint Force Quarterly* 63 (4th Quarter 2011), <http://www.ndu.edu/press/why-iran-didnt-admit-stuxnet.html>. In the wake of Stuxnet, one Iranian official noted that "[a]n electronic war has been launched against Iran," but there was never an official government statement endorsing that view. Atul Aneja, "Under Cyber-Attack, Says Iran," *Hindu*, 26 September 2010, <http://www.thehindu.com/news/international/article797363.ece>.
35. CCDCOE, *Cyber Attacks against Georgia*, 4.
36. Quoting Henry Lewis Stimson, secretary of state under Herbert Hoover, justifying closing the "Black Chamber" in 1929, the code-breaking office. Documentation of espionage dates back

thousands of years. Egypt had an organized intelligence service 5,000 years ago, and espionage is one of the dominant themes in Sun Tzu's *Art of War* 2,500 years ago. Kurt D. Singer, *Three Thousand Years of Espionage* (New York: Books for Libraries Press, 1948), vii.

37. Some legal scholars argue that espionage is a violation of sovereignty, but this is the minority view. See Manuel R. Garcia-Mora, "Treason, Sedition and Espionage as Political Offenses under the Law of Extradition," *University of Pittsburgh Law Review* 26, no. 65 (1964): 79–80; and Quincy Wright, "Espionage and the Doctrine of Non-Intervention in Internal Affairs," in *Essays on Espionage and International Law*, ed. Roland J. Stranger (Columbus: Ohio State University Press, 1962), 12. See 18 U.S.C., pt. 1, chap. 37, "Espionage and Censorship," and 18 U.S.C., §§ 793–98, for the US domestic law.

38. For example, in July 2010 the United States and Russia exchanged spies after the FBI uncovered a Russian sleeper cell. See "U.S. Confirms Successful Exchange of Spies," *CBS News*, 9 July 2010, <http://www.cbsnews.com/stories/2010/07/09/world/main6661165.shtml>.

39. CDR Roger Scott, "Territorially Intrusive Intelligence Collection and International Law," *Air Force Law Review* 46 (1999): 217–18.

40. Josh Zachry, associate director for research operations, Institute for Cybersecurity, University of San Antonio, quoted in "Cyber Espionage Threatens Global Security," *Intelligencesearch.com*, <http://www.intelligencesearch.com/ia158.html>.

41. Tom Gjelten, "Cyber Insecurity: U.S. Struggles to Confront Threat," *NPR.org*, <http://www.npr.org/templates/story/story.php?storyId=125578576>.

42. See discussion of "effects" below.

43. Prescott Winter, "Cybersecurity—Governments Need to Cooperate," *Cyber Threat* blog, 8 April 2010, <http://blogs.computerworlduk.com/cyber-threat/2010/04/cybersecurity--governments-need-to-cooperate/index.htm#>.

44. A notable exception is the case of mistake of fact or accident, such as air strikes that hit the wrong targets or targets that were unintentionally mischaracterized, in which case the victim state and the international community may assess the reasonableness of the mistake before characterizing the action under the law of war. See Daniel Williams, "NATO Missiles Hit Chinese Embassy," *Washington Post*, 8 May 1999, A-1; and "US Warplanes 'Bomb Afghan Wedding Party,'" *Independent*, 6 November 2008.

45. Art. 2(4) of the UN Charter prohibits even threats of a use of force. As states have proven themselves unwilling to give up espionage, it is unlikely the "threat of force" prohibition will be given a broad interpretation in the case of cyber activities. This might mean that states will be free under international law to implant dual-use computer code and be poised to strike, while defending states would legally be expected to wait until the moment the code was converted before acting in self-defense. A fuller discussion of this interesting issue is beyond the scope of this article.

46. *Mens rea* is a legal term referring to the intent element necessary to be convicted of a crime.

47. In a 2007 Department of Homeland Security exercise called Aurora, controlled hacking into a replica of a power plant control system enabled researchers to change the operation of a generator, resulting in its violent physical destruction. "Staged Cyber Attack Reveals Vulnerability in Power Grid," *CNN*, 26 September 2007, http://articles.cnn.com/2007-09-26/us/power.at.risk_1_generator-cyber-attack-electric-infrastructure?_s=PM:US.

48. Other factors have been suggested to form a test for a use of force. The most commonly cited is Prof. Mike Schmitt's six-part test for cyber attack, which requires assessing cyber actions for severity, immediacy, directness, invasiveness, measurability, and presumptive legitimacy. Although this is a rational test for analyzing cyber actions post facto, we would argue that only the first—severity—is necessary to determine if the event qualifies as an attack. The lightning speed

of cyber actions makes swift decision making critical, and it is unlikely nations will have the information or the time to consider these factors in the heat of potential battle. Professor Schmitt's test could be very useful in determining whether a cyber action violated an international norm not predicated on a use of force, such as the principle of nonintervention. See Michael N. Schmitt, "Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework," *Columbia Journal of Transnational Law* 37 (1998–99): 885; and *The Principle of Non-Intervention in Contemporary International Law: Non-Interference in a State's Internal Affairs Used to Be a Rule of International Law: Is It Still?*, Chatham House discussion group summary, http://www.chathamhouse.org.uk/files/6567_il280207.pdf.

49. John Markoff and Thom Shanker, "Halted '03 Iraq Plan Illustrates U.S. Fear of Cyberwar Risk," *New York Times*, 1 August 2009.

50. Financial Inquiry Commission, *Final Report of the National Commission of the Causes of the Financial and Economic Crisis in the United States*, January 2011, <http://www.fcic.gov/report>.

51. Devlin Barrett, Jenny Strasburg, and Jacob Bunge, "NASDAQ Confirms a Breach in Network," *Wall Street Journal*, 7 February 2011. For a general discussion of the National Association of Securities Dealers Automated Quotation (NASDAQ), see "NASDAQ Wiki," *Motley Fool*, <http://wiki.fool.com/Nasdaq>.

52. See "U.S.-Soviet/Russian Arms Control," *Arms Control Today*, June 2002, http://www.armscontrol.org/act/2002_06/factfilejune02.

53. *International Strategy for Cyberspace: Prosperity, Security in a Networked World* (Washington: White House, May 2011), 9.

54. Gerry Shih, "Twitter to Restrict User Content in Some Countries," *Reuters*, 27 January 2012, <http://in.reuters.com/article/2012/01/26/twitter-idINDEE80P0IR20120126>.

55. *International Strategy for Cyberspace*, 12; and *Department of Defense Strategy for Operating in Cyberspace* (Washington: DoD, July 2011), 10.

Book Reviews

Critical Code: Software Producibility for Defense by the National Research Council. National Academies Press, 2010, 160 pp., \$34.75.

As the title implies, computer software is critical to the mission of the Department of Defense. Hence, the National Research Council (NRC) was tasked by the Office of the Secretary of Defense in 2006 to analyze and make recommendation on all aspects of software development and sustainability pertaining to the DoD. Specifically, the NRC's Committee for Advancing Software Intensive Systems Producibility delved into current DoD processes for building or buying software for the vast number of systems—both weapon-related and administrative—used by the military.

The research group, comprised of 14 recognized software gurus from corporate America and renowned American universities, complied with the OSD's direction to analyze current software producibility in the military and to debunk myths surrounding defense software. Software producibility, as defined by the NRC, is "the capacity to design, produce, assure, and evolve innovative software-intensive systems in a predictable manner while effectively managing risk, cost, schedule, and complexity."

The committee outlined its findings associated with eight software producibility myths and made salient recommendations on what the DoD should do to fix the problems. These myths range from software producibility challenges associated with management and processes to the one that "There is sufficient software research already underway, sponsored primarily by NSF [National Science Foundation] and other basic science agencies to meet the DoD's software needs."

In the first chapter, the committee defines the role of software in the defense industry and how the DoD addresses the software needs of the military. One statistic in this chapter readily defines the salient issue of the study—"the percentage of system functions performed by software has risen from 8 percent in the F-4 in 1960, to 45 percent of the F-16 in 1982, to 80 percent of the F-22 in 2000." This is the fundamental reason why the DoD must examine its current way of handling military software to ensure it is properly designed and implemented. According to the study, the DoD has actually decreased its software producibility in recent years by contracting out production or purchasing "off the shelf" from Microsoft and other software companies, both US and foreign.

Chapter 2 addresses a common belief in the military that commercial off-the-shelf (COTS) software avoids the huge costs of training military members to write software and produce it in-house. It also discusses risks associated with software and how the DoD can manage such risks. The DoD cannot afford to lag behind in this area because its adversaries are constantly looking at new ways to hack into

computer systems and writing counter-cyber programs to protect their own IT infrastructures. The only way to beat the threat and accept risks is to “engage experts outside the DoD” to be effective and stay ahead in software producibility.

Another issue identified by the NRC is that software is not at a plateau but is growing along with the technology surrounding it. Moore’s Law is alive and well in the software realm and not just relevant to firmware or hardware.

The NRC study stresses the importance of leaders getting involved in the architecture side of software. According to the committee, “Architecture is an important enabler of reuse and the key to system evolution, enabling management of future uncertainty.” It is something that must be managed not only during the first phases of development and employment, but all the way through the life of the system using that software. The DoD must learn from corporate America and use those lessons to aid in mitigating the risks in architecture systems.

Chapter 4 discusses the importance of quality assurance in software, both the defense and civilian systems that the DoD uses. Without quality assurance, all the systems that use millions of lines of code could put the operator in harm’s way and/or cost billions of dollars to fix. Weak software is also a breeding ground for cyber attacks and infiltration by the enemy. Studies suggest that “overall software assurance costs account for 30 to 50 percent of the total project costs for most software projects.” Software assurance is not an inexpensive endeavor but one that must be incorporated at the start of the software life cycle.

The final chapter “summarizes and recommends technology research areas as critical to the advancement of defense software producibility.” The laundry list varies from DoD influence on academic research and development; to the impact of past investments, challenges, and opportunities for investment; to areas for future research investment.

Overall, the message in *Critical Code* is very pertinent to today’s interest in cyber security and the software that the DoD uses in ensuring national security. The text is rather technical and a bit hard to follow at times. However, it is a good read for cyber officers and leaders (both military and civilian) to ensure nothing slips through the cracks and causes catastrophic issues with the myriad of systems in the DoD.

Lt Col Deborah Dusek, USAF
Offutt AFB, Nebraska

Airpower for Strategic Effect by Colin S. Gray. Air University Press, Air Force Research Institute, 2012, 367 pp., available free at http://aupress.au.af.mil/digital/pdf/book/b_122_Airpower.pdf. Commercial version published by Columbia/Hurst, 2012, 288 pp., \$55.00.

In this expansive assessment of airpower’s steady rise in salience from its fledgling days to today’s combat involvements, Colin Gray, a prolific strategist of long-standing scholarly reputation, has produced an outstanding tutorial for Airmen by addressing the air weapon in the context of what he calls its abiding “strategic

narrative” (p. 1). His book is not about the tangibles of airpower—the platforms, munitions, and associated support systems—that make up its hardware ingredients. Rather, it is about how one should think about airpower’s larger meaning and significance.

This important new book begs to be read by airpower’s doers as well as its thinkers—and at all rank and command levels. In explaining why, Gray notes that his intent in writing it was “to contribute to a better strategic understanding of airpower to improve the *practice* of airpower” (p. 2; emphasis added). Toward that end, he stresses that his purpose was not to indulge in debate over air doctrine but “to help sharpen the ability of readers themselves to engage in such debate” (p. 4)—most notably in the all-important policy arena in which the most intractable cross-service disagreements over roles and resources get adjudicated.

Gray’s central theme is that airpower generates strategic effect. More to the point, he maintains, it is a tactical equity that operates—ideally—with strategic consequences. To him, “strategic” does not inhere in the equity’s physical characteristics, such as an aircraft’s range or payload, but in what it can do by way of producing desired results. From his perspective, a strategic effect is, first and foremost, that which enables outcome-determining results. And producing such results is quintessentially the stock in trade of American airpower as it has progressively evolved since Vietnam.

With this unifying principle as his point of departure, Gray improves on Brig Gen William “Billy” Mitchell’s definition of airpower by characterizing it more helpfully as “the ability to do something [*strategically useful*] in the air” (p. 9; emphasis in original). He further stresses—as his book’s title well reflects—that only by producing desired *effects* can airpower’s use in warfare be deemed successful.

In addressing the predominance of today’s low-intensity insurgent challenges, in which kinetic air attacks have largely been overshadowed by ground forces in the starring role, Gray takes a long view of airpower’s relevance and potential by appraising the air weapon in the broader context in which its payoff will ultimately be registered. His survey of airpower’s combat use over time shows convincingly how the relative importance of the air weapon is neither universal nor unchanging but totally dependent on the circumstances of a confrontation.

More to the point here, when viewed operationally, airpower can be everything from single-handedly decisive to wholly supportive of a combatant commander’s needs. Because its relative import, like that of all other force elements, hinges directly on how its comparative advantages relate to a commander’s most immediate concerns, Gray reminds us that airpower need not disappoint when it is not the main producer of desired outcomes. Indeed, he rightly notes, the notion that airpower should be able to perform effectively in all forms of combat unaided by other force elements is both an absurd measure of its value and a baseless arguing point. By misguidedly espousing this point over many decades, airpower’s most outspoken advocates have done their cause a major disservice.

It naturally follows from this, Gray adds, that whenever airpower has been said to have “failed,” it has only been because more was expected of it than it could deliver. After all, *any* tool can appear deficient if used unwisely or irresponsibly. In this regard, Gray notes how a long history of overpromising on the part of airpower’s most vocal proponents has needlessly sold the air weapon short for what it is actually able to deliver to joint force commanders today—and not just in high-intensity combat but in *all* forms of operations across the conflict spectrum.

To be sure, Airmen of action may find it trying at times to remain patient with Gray’s always purposeful but also often discursive walk through the intellectual thickets of airpower theory. In a frank admission of his own appreciation of those readers who will be all too eager for him to get to his point, Gray freely concedes how “theory and theorists often are regarded with disdain by the people ‘out there, doing it,’ when in truth the purpose of the theory enterprise is both to reduce the risks to the warriors and to help make their efforts more useful vis-à-vis the operational goals that are set” (p. 41).

Yet were there ever an instance in which patience should have its rewards for mission-oriented Airmen of action, it is plainly here, for *Airpower for Strategic Effect* offers an uncommonly thoughtful application of informed intellect to an explanation of how modern air warfare capabilities should be understood. In his last chapter, Gray underscores in this regard the important truth that “airpower theory helps educate airpower strategists,” rightly calling it “theory for practice” (p. 275). Furthermore, he instructively adds, it “educates those who write airpower doctrine and serves as a filter against dangerous viruses” (p. 276).

At bottom, the purpose of Gray’s treatise is not to extol airpower but to make coherent sense of it by providing informed insights into it and about it that are timeless. For Airmen of all ranks, the greatest value that its appreciation of the air weapon can offer is to help them think more reflectively about their calling and to articulate its foundational principles more effectively in the councils of war planning. For woven throughout the book is a compelling explication of what modern airpower entails in its most inner strategic essence. The ultimate aim of that explication is to improve the real-world *practice* of airpower by operators at all levels most responsible for its effective use.

Benjamin S. Lambeth, PhD

*Senior Fellow, Center for Strategic and Budgetary Assessments
Washington, DC*

Chinese Aerospace Power: Evolving Maritime Roles edited by Andrew S. Erickson and Lyle J. Goldstein. Naval Institute Press, 2011, 544 pp., \$52.95.

Andrew Erickson and Lyle Goldstein, two prominent China scholars at the Naval War College, fill an important interdisciplinary niche with this book by bringing together an all-star team of authors from both the Air Force and the Navy communities. By no means a light read, *Chinese Aerospace Power* is in fact a compendium, a compilation of 27 essays authored by an illustrious group including

admirals, intelligence analysts, private-sector experts, and former defense attachés. The fifth volume in a series on Chinese military developments in the maritime arena, the book stands as a stark reminder that China's rise, while impressive to date, can only be expected to accelerate in coming decades.

Due to the diversity of authors and the range of topics covered, the book does not support any single, overarching thesis. If there is one recurring theme, however, it might be this: Chinese military power is rapidly increasing, and American primacy in the Pacific is threatened as a result. Changes in the balance of aerospace power over China's littoral waters have far-reaching strategic consequences for American policymakers. This book explains both how and why—in dense detail.

While overall a fascinating read for anyone with a strong interest and/or background in Chinese military affairs, one difficulty with the book stems from the sheer scale of the undertaking. At times the reader is left in something of a fog, having to piece together enormous amounts of highly technical information—a bit like being shown a sky full of stars but no constellations. Admittedly, this is a difficulty common to multiple-author works, whereas authors writing alone or in small teams have the ability to lace a clear thesis throughout even the most complex subject matter. As a result, some information is repeated in *Chinese Aerospace Power* a bit more than one would like.

Nevertheless, for those who find the technical, even obscure details interesting (this reviewer included), this book is a real treasure trove. The work spans six broad subject areas, each of which has been the subject of considerable literature in recent years: the emerging roles of Chinese aerospace power; the intelligence, surveillance, and reconnaissance (ISR) and counter-ISR capabilities of the People's Liberation Army (PLA); PLA aerospace strategy; air-launched cruise missiles; ballistic missiles; and the implications of Chinese aerospace power for the US military. Strategic studies aficionados will find the chapters on strategy and missile development particularly worthwhile.

Several authors explain how, properly coordinated, Chinese aerospace power has the potential to vastly enhance antiaccess capacity, pushing foreign forces away from Chinese shores and affording the PLA the strategic depth to turn its energies toward other concerns, such as the "active" side of its doctrine of "active defense." Paul Giarra, Andrew Erickson, and David Yang excel in addressing one of the key components of China's emerging antiaccess capacity: antiship ballistic missiles (ASBM), which RADM Eric McVadon, USN, retired, has elsewhere argued could have implications similar to those of China's first successful nuclear test in 1964 (he reasserts this position in the book's final chapter). As several authors persuasively argue, if the PLA can deploy ASBMs capable of hitting moving carrier strike groups (CSG), US Navy power projection calculations in the region could be "upended." For decades, the heart and soul of US Navy power *Forward . . . From the Sea* has been the aircraft carrier, in large part because it could move with relative impunity on the high seas. American carriers, for example, deployed to the Taiwan Strait in 1995 and 1996 as a show of force in defense of Taiwanese democracy; until now,

the Chinese government has been unable to counter such a threat. Several authors make a compelling case that this could change in a matter of just years.

Discussion of PLA aircraft development likewise gives one cause for concern. Pushing the US Navy away from Chinese shores could give the PLA the operational breathing room needed to achieve air superiority over Taiwan. Chapters on PLA Air Force (PLAAF) power share a theme with the ASBM chapters discussed above: the balance is tilting in China's favor. Fourth-generation fighters now make up approximately 20–25 percent of the 2,000-plus combat aircraft in the PLA arsenal, and that ratio is expected to approach 50 percent in the coming decade. Backed by the bristling missile defense of the Chinese Second Artillery Corps, Chinese air superiority over Taiwan could be achieved in short order.

One of the more concerning takeaways from this book is the limited set of options available to American policymakers. To preserve the balance in America's favor would be enormously—even prohibitively—expensive. Maintaining a safe distance off China's coast could soon mean short-range aircraft in US Navy air wings could have little real utility, cruise missiles could lose their efficacy, and Marine Corps amphibious landings “would not be realistic.” Refitting the US fleet would come at enormous cost, which is why in the final chapter, Admiral McVadon argues that the benefits of Sino-American cooperation could soon outweigh the costs. The ultimate takeaway might therefore be this: the era of “rising China” may fast be coming to an end—China is on the verge of being fully risen.

Imagine a world 10 years from now. China's growing battery of nuclear ICBMs has the capacity to reach all corners of the continental United States. American Pacific island bases and CSGs once offering protection to Taiwan now sit within range of a devastatingly large stockpile of missiles in mainland China. Kadena AFB in Japan begins each day confronted by the bleak fact that it could be grounded for a week or more by a Chinese first strike. Fourth-generation PLAAF fighters are on standby, ready to disrupt US efforts to gain air superiority should armed combat erupt near Chinese shores. Any effort to deploy American fourth- and fifth-generation fighters over the Chinese mainland means subjecting them to the world's most fearsome surface-to-air missile force. In short, Americans are vulnerable at home, the ability of the US military to assert control of the Pacific theater is greatly compromised, and American retaliatory options are limited mostly to long-range missile and bomber strikes. China is now a fortress. At this point, China announces its new grand strategy: the deployment of carrier groups capable of circling the globe. China's power surge accelerates.

Anyone who finds such a future difficult to imagine would benefit from reading this book. Not only is such a future imaginable, those who read *Chinese Aerospace Power* may very well come to expect it.

Capt Paul A. Stempel, USAF
Joint Base Andrews, Maryland



CYBER CONFERENCE AT MAXWELL AFB

CYBER POWER

The Quest Toward Common Ground

Cyber Power, National Security and Collective Action in Cyberspace

The Air Force Research Institute (AFRI) at Air University is conducting a symposium series to contribute to a better understanding of the structural sources of cybersecurity challenges. Our project considers current and potential ways to strengthen and expand the way we are organizing collective responses to cyber vulnerabilities and threats. Conference breakout sessions will stimulate and develop experientially informed, interdisciplinary research in an effort to generate rapid insights about key topic areas that may not be well understood.

10 - 11 October 2012
Maxwell AFB
Montgomery, Alabama



For conference registration and more, go to:

Public Website: <http://afri.au.af.mil/cyber>

Twitter: <http://www.twitter.com/afri09> #afri09



**Georgia
Tech** | **Research
Institute**





"Aim High . . . Fly-Fight-Win"

