Keeping Cool

Steps for Avoiding Conflict and Escalation in Cyberspace

Roger Hurwitz

International cyber conflicts will become more frequent and intense. Several current trends prompt this pessimism: ongoing quantitative and qualitative cyber arms races among state actors, the proliferation of cyber weapons among state and non-state actors, a widely shared conventional wisdom that these weapons favor offense over defense, and no broadly accepted definition of the cyber attacks that amount to the use of force. In 2010, a United Nations group of government experts noted that the absence of redlines in cyberspace could foster miscalculations, misperceptions and, in turn, unwanted conflicts. With the rare agreement of the Chinese, Russian, and U.S. members, it issued a call, endorsed by the General Assembly, for states to discuss norms of behavior in cyberspace.¹

While this action did not lead to any agreements on the contents and applications for such norms, the declarations on this matter by some states have clarified the ideological fissures that divide them. Their disputes over terminology and extent have also blocked adoptions at the regional level of less demanding confidence-building measures for cyber (iCBMs), which could reduce states’ mutual suspicions.² Meanwhile, over twenty countries have created dedicated
military cyber commands with responsibilities that probably include targeting and operations in other states’ networks, like those that U.S. Presidential Directive 20 assigns the U.S. Cyber Command. Given these trends, states are unlikely to stop halfway in cyber conflicts or in using cyber weapons in the context of kinetic wars.

**The Coming Heat.** The concern of this essay is the politically feasible steps that states might take to avoid cyber escalation and the spill over of cyber conflicts into the kinetic realm. To the extent that policy and decision makers may share this concern, their motivation to avoid these interactions will be different and weaker than those they have regarding protracted or escalating kinetic conflicts. The consequences that can be anticipated for such wars, especially among nuclear capable parties, typically bias decision makers toward conflict avoidance. They may use force as a last resort, and only when convinced their conventional military capabilities can achieve decisive and quick advantage over the adversary.

These leaders can be more sanguine about the effects of cyber conflicts, notwithstanding their occasionally expressed anxieties about attacks on critical infrastructures. Cyber weapons alone will not produce the devastation of kinetic conflicts, much less of a nuclear war. While distributed denials of service (DDoS) or damaging attacks can cause significant disruption in developed countries, where societies and economies depend on digitally based, highly coupled infrastructure, the redundancy and resilience in the infrastructure can limit the effects of these attacks to tolerable levels. In any case, the costs of disasters, which are analogues or precedents for these attacks, such as the 2003 blackout in the United States and Canada or the 2007 DDoS in Estonia, tend to be significantly less than those of hurricanes and other natural disasters.

More worrisome cyber threats to national security include the possibility of a “cyber Pearl Harbor” – approximated in Georgia 2008 – that would disable a country’s command, control and intelligence capabilities at the start of a kinetic war, or, in the case of the United States and other technologically advanced countries, the cyber enabled theft of their intellectual property on a massive scale. However, as chatter about pre-emption, “cyber operations,” and “hack backs” suggest, these threats have made leaders and planners tilt toward strategies that involve rather than avoid escalation – that “raise the stakes” and “keep one’s options open.”

Nevertheless, leaders seem to have some interest in not making cyberspace a battlefield à l’outrance, in view the...
inevitable degradation of subsequent peacetime use that would lead to more securitization, more fragmentation, less economic and social development based on the Internet, loss of a platform for discussing global problems, greater difficulty in projecting soft power and in using cyber to penetrate the secrets of foreign governments. At least, most leaders agree on the economic blessings due to the Internet, and intergovernmental meetings on cyber security often start with speakers recounting such golden eggs, as if to remind themselves not to kill the goose.

This pale motivation may be further diminished by decisions to avoid conflict having to be made in a more complex and uncertain contexts than those found for conventional dyadic clashes and potential nuclear ones. When only the United States and USSR were nuclear-ready and confident of their second strike capabilities, they could imagine their potentially hostile interactions as repeated games of “Prisoner’s Dilemma” or “Chicken,” where cooperative moves – conflict avoidance or de-escalation – could be rationalized as being the best self-interested choices. This stark mental landscape enabled convergence of the dyad’s beliefs on the risk and handling of nuclear weapons, even before they discussed these matters directly. Their learning in turn set precedents for other nations that acquired nuclear weapons.

Today’s cyber environment, however, is cluttered and destined to become more so, with numerous state and non-state actors having offensive cyber weapons. Hence attribution of an attack is problematic enough to undermine the credibility of a threat of retaliation intended to deter an attack. Several bad solutions to this problem have been suggested, such as threatening disproportional retaliation when a cyber attack can be attributed with certainty to an actor that is believed to be responsible for previous, less attributable attacks. Such moves can be anticipated to generate enough new risks for the would-be retaliator as to raise doubts it would follow through on them.

Furthermore, a game theoretic analysis of choices in this environment verges on the intractable. For more than two players, with each having multiple modes of cyber attacks ranging from network exploitation to damaging critical infrastructure, and several strategies for avoiding conflict or escalation, a particular strategy to avoid conflict does not emerge as the strong self-interested choice, if it emerges at all. This failure can have several different causes. First, there are likely to be multiple equilibriums or combinations of strategies from which the respective players cannot unilaterally depart without risking loss, even though the payoffs at another equilibrium are better. For example, states might avoid conflict by disconnecting their national networks from the Internet, but that strategy foregoes the benefits that accrue from being part of a global network. States, which adopted this strategy but later wanted to reconnect would need to reassure one another their new strategies would not be exploited. The situation is analogous to that of two players who are locked into mutual defection in repeated Prisoner’s Dilemma games because each can afford a change to cooperation only if they change together. Second, a general condition of conflict avoidance
or peace in cyberspace can be considered a public good, and it is subject to exploitation by players who believe they can practice undetected a conflictive strategy, such as cyber espionage, without risking loss of the public good’s benefits. Finally, given the possibilities of misattribution and misperception in this environment, a player’s best strategy might be to spoof an attack by a second player on a third, with the aim of precipitating a conflict between the two that will weaken both of them.

Given states’ limited motivations for avoiding cyber conflicts, guidelines and practices that simplify their choice environments rather than seek to build trust among states will have better chances in keeping cyberspace relatively stable. These will reduce ambiguities in how states should regard and respond to various cyber threats, and differentiate responses according to the threat involved. They are a minimalist set, which states might find expedient to accept and whose practice will contribute to the stability of cyberspace. They are not intended to create an institutional or legal order there.

Reducing Ambiguities. States should set some broadly accepted redlines in cyberspace. States have not yet agreed to designate some types of attacks as a “use of force” in violation of the United Nations charter’s article 2 (4) and to further characterize a subset of these as “armed attacks,” per article 51, to which the attacked party may respond with force. One obstacle to agreements may be the cognitive difficulty of thinking of cyber attacks as a “use of force,” since they are conducted by non-kinetic means such as malware and their damage (including damage as a consequence of disruption) is caused indirectly. As such, they raise again the question whether psychological, economic, and other non-military actions against states constitute use of force or fall into another category, which would preclude their rising to “armed attacks.”

Another obstacle may be a fear among some states, including Russia and China, that by recognizing certain cyber actions as casus belli, they would compromise their earlier calls for demilitarization of cyberspace made perhaps in the hope of freezing the United States’ advantage in the development of cyber weapons. For its part, the United States has believed that designation of redlines would be tantamount to a declaratory policy that could reduce its options in the event of an attack. Nevertheless, states increasingly recognize that cyberspace is militarized, and seem to be converging in both global and bilateral discussions on the applicability of the laws of war to it — both jus ad bellum and jus in bello — as part of the applicability of international law to cyberspace.

Even so, differences among states as to which cyber actions are “use of force” or “armed attacks” could prevent unambiguous agreements. Russia and China, per their Code of Conduct tabled at the UN in September 2011, would seem to include in “use of force” the distribution by digital networks from state A to state B of information that state B’s government considers destabilizing of its internal affairs. Western liberal democracies reject this approach, considering it a demand that a host country stifle exiles’ political
States should publicly rank types of cyber exploits and attacks according to their severity. Escalations in conflicts are ordinarily caused by reciprocal, disproportional responses to provocations. That is, state B over-reacts to state A’s action, then state A over-reacts to state B’s over-reaction, and so forth. Hence proportionality, as required by the Law of Armed Conflict, is a means of stabilizing international order as well as containing cruelty. However, what constitutes a proportional response to a cyber operation may well lie in the eye of the beholder, particularly when there are several types of operations -- making proportionality more than a matter of quantity -- but traditions in their use are lacking.

This recommendation does not anticipate that states will agree on a single scale, or that all states will publish their scales if their cyber security planners have developed one. To the extent respective scales are known, one state considering a cyber action against another can better estimate how its target would evaluate that action, and, reciprocally, if the attack occurs, the target can better understand the presumed attacker’s intentions. The point is that an actor in a dyadic conflict who takes into account the other actor’s

By informing each other of how they respectively rate various cyber operations, China and the United States can better understand the intentions behind and expected responses to the other’s actions.

unrest. These more ambiguous cases could be argued before the Security Council or other appropriate forums as or after they occur. In these instances, the argument would probably be based on specific effects (e.g., scope, duration, lethality, pain inflicted) rather than their means. For example, while most DDoS attacks are disruptive, annoying, and inconveniencing rather than damaging, some, notably the attacks on Estonia in 2007, have been claimed to rise to “armed attack,” given their invasiveness and frustration of societal life.¹⁵

speech if their home country’s government so requests.

The ideological impasse could be partly resolved by respective sides recognizing they will not convince their opponents, but that almost all can publicly agree that other cyber actions are “uses of force” or “armed attacks” -- the most obvious being direct, cybereaged sabotage of critical infrastructure resulting in damage to property and possibly loss of lives, with precedent in the attack on Iranian centrifuges.¹⁴ Any agreement on a minimal set would need to provide that the status of other attacks, such as disruptions or incitements to civil unrest. These more ambiguous cases could be argued before the Security Council or other appropriate forums as or after they occur. In these instances, the argument would probably be based on specific effects (e.g., scope, duration, lethality, pain inflicted) rather than their means. For example, while most DDoS attacks are disruptive, annoying, and inconveniencing rather than damaging, some, notably the attacks on Estonia in 2007, have been claimed to rise to “armed attack,” given their invasiveness and frustration of societal life.¹⁵
perceptions can act more proportionately and prudently. Without information or respect for these perceptions, spirals of misestimates, misunderstandings, irritation and escalation are likely. Consider the conflict in 2013 between the United States and China over cybered industrial espionage against American firms, which the US alleges is sponsored by the Chinese government. U.S. leaders have condemned it as "economic warfare," "the greatest theft of wealth in history," and, at the very least, a violation of international law. When they are not denying responsibility and claiming they are also victims of such exploits, Chinese leaders have spoken of their country’s suffering a “century of humiliation” by colonial powers, perhaps implying that taking intellectual property from Western firms is appropriate compensation. They could add that traditional Chinese culture emphasizes copying works of excellence, but places less value than Western culture on originality and intellectual property rights.

On these views, the U.S. overrates China’s intentions, while China underestimates how seriously the U.S. regards the espionage. Thus China would likely regard as disproportional an aggressive cybered U.S. response, such as hacking the networks of the Chinese military units that conduct espionage, or cross-domain retaliation, such as trade sanctions. On the other hand, unless China does begin to rein in the espionage, the United States will feel compelled, given its perceptions, to take such action. By informing each other of how they respectively rate various cyber operations, China and the United States can better understand the intentions behind and expected responses to the other’s actions. Their exchange might lead to convergence in the ratings, prompting some adjustments in respective cyber behavior and planning, and providing a basis for mutually defined proportional responses and effective signaling in future cyber conflicts, while avoiding the antagonism of making and rejecting formal demands. The U.S. and China appear to have taken a step in this direction by creating a bilateral cyber security working group in the summer of 2013.

Another helpful step is the publication of a semi-official U.S. Cyber Command view of the relative severity of cyber operations, illustrated by actual and foreseeable cyber incidents. Dubbed a "spectrum of operations," it has fifteen levels, based mainly on effects, ranging from probes and unauthorized access to another state’s networks, through disruptions of those networks unto operations that rise to the use of force and armed attacks. These are further characterized by their stealth, with access operations being very stealthy, while cyber attacks much less so, and the responses that the various operations might elicit.

Targets of exploits and disruptions may respond with counter measures such as disruptions of their own making, but are unlikely to retaliate with force; attacks which destroy official government systems, damage property or cause injury or loss of life can be expected to provoke use of force in return. The Stuxnet attack, described as damaging 1000 Iranian centrifuges, is cited for operations causing damage or injury, but a more severe level – operations causing death or serious injury
is specified by the use of cyber in the context of kinetic conflicts to destroy command and control networks, fuel storage, aircraft and ships through subverting their digital controls.

Levels of degradation, disruption, and damage of networks are distinguished according to ownership of the targets (i.e., private vs. government, and how critical they are). These distinctions align with the notions of invasiveness and frustration of societal life mentioned earlier, but they are not categories of international law. Convergence on a scale among a government’s agencies or a group of states will be a political agreement rather than a result of legal arguments, and be informed as much by policy needs and elite opinions as by international and domestic laws. Thus, Washington policy makers, given their rhetoric in the summer of 2013, might prefer a new version of the spectrum that distinguished industrial espionage and upgraded its severity and provocation, rather than one which seems to follow the ambivalence of traditional international law to espionage.

It is important to also note that a spectrum of cyber operations aligns with only part of the possible ranged actions in international conflicts, which can vary from the slightly conflictive act of one state urging a particular policy on another to severe military assaults or clashes with prospective casualties several orders of magnitude greater than can be imagined for any cyber attacks. On this view, cyber conflict may be preferred to other types, providing it does not spill over into the kinetic conflict. The next sections suggest a few ways of reducing that likelihood of that happening.

**Differentiating Responses to Cyber Provocations.** States should enhance their passive cyber defense, resilience, and recovery capabilities to increase the possibilities of deterrence through denial, and also reduce recourse to active defense (i.e., cyber operations in a presumed attacker’s networks). Although the present technological balance favors offense over defense, potential victims can force exploiters and attackers to work harder by making policy level changes—some able to be implemented by individual organizations, some requiring national legislation. These include organizations’ educating their personnel to the dangers of social engineering, automated patching of operating systems, acquiring and sharing threat information on a voluntary or mandated basis, quarantining compromised systems, detection and remediation of (botnet) captured computers, and improved supply chain assurance through liability and insurance leverage on leading vendors.

Resilience of critical information systems can be enhanced through redundant bandwidth and storage, and ephemeral virtualized networks in which nodes can be added or excluded according to security policies. These moves are facilitated by cloud computing. That said, the attacks on American and European financial institutions in the fall of 2012 demonstrated that the cloud servers could also increase the challenge (i.e., the bandwidth) of DDoS attacks by an order or more of magnitude.

To the extent such measures discourage attackers and reduce their suc-
cess, they also reduce recourse to more risky active defense. These risks include cyber operations against an innocent third party because the attack was misattributed, and consequences well beyond suppression of the attacker’s servers because of the configuration of the attacker’s networks. Such cases would widen the circle of provocations and grievance, and, especially if parties’ critical systems are impaired, they might introduce a sense of desperation that nudges the conflict closer to use of kinetic weapons.

The anticipation that active defense may trigger escalations is a reason that some analysts have proposed a norm of reversibility in cyber exploits and attacks. For example, an attack on hard drives would encrypt rather than erase their files or physically damage them so that an adversary could not access their information during conflict, but could recover it afterward when the attacker decrypted them. Their proposal is consistent with contemporary concepts of war fighting, which prize sapping the adversary’s capability and will to fight over its annihilation.

In the framework of these concepts, associated with the “Revolution in Military Affairs,” and given the enormous dependence of modern governments and military on digital networks, cyber operations seem the weapons of choice par excellent. Yet, one may doubt that cyber conflicts will be conducted with as little emotion as these proposals foresee, and consequently steps that minimize recourse to cyber operations would be more prudent than those seeking to ritualize them.

States should delay responding to attacks if time is needed to confirm attribution or craft a suitable response. Nobel laureate and game theorist Thomas Schelling pointed out that escalations occur when people in conflict do not stop to think, by which he meant that a rushed response to an attack will likely be disproportionate and provoke a larger attack in the next round. Moreover, because “false flag” cyber attacks are easily mounted, initial attributions for them are more likely to be mistaken than those for kinetic attacks so that a rushed response risks an undesired widening of the conflict without deterring the attacker. However, in the context of an intense conflict or hot war when a state’s leaders believe their digital resources and systems are threatened, their inclination will be to respond quickly and preemptively to suppress the potential attacks unless they are

Cyberspace is under-institutionalized, under-policed, and lacking normative traditions, but it is not an isolated, Hobbesian state of nature where conflict among states is settled only by force.
also confident in their cyber defenses. Thus, the pause before responding that such defense can afford is another reason to invest in them.

The problem of how and when to respond can be subject to a cost-benefit analysis that should take into account the costs of acting on the basis of misattribution and the risks of escalation – Robert Axelrod’s “shadow of the future” – as well as the costs of further attacks, on one hand, and the probability of deterring them, on the other. A state’s delay in responding in order for computer forensics to produce better attributions does not necessarily diminish a state’s power to deter. On the contrary, subsequent, proportional retaliation based on evidence for the attribution that is shared with other states can enhance the state’s reputation for acting forcibly and justly in defending its interests.

**States should seek responses to exploits and attacks at forums and in measures outside of cyberspace.** Cyberspace is under-institutionalized, under-policed, and lacking normative traditions, but it is not an isolated, Hobbesian state of nature where conflict among states is settled only by force. Many diplomatic, political, economic, and cultural ties connect states, and a state can leverage some for support against a cyber attack or to find an appropriate response to an attacker.

While NATO deemed that the 2007 DDoS attacks on member state Estonia were not a matter for collective security, some member states did provide computer resources and diplomatic support that enabled Estonia to withstand the attacks and perhaps hastened their end. Additionally, Estonia sought afterwards through legal channels to identify and apprehend individuals in Russia responsible for the attacks. The Russian prosecutors’ refusal to cooperate on grounds that cyber crimes were not covered by the mutual assistance treaty between Russia and Estonia should not suggest that seeking legal recourse is worthless, but rather that states must endeavor to expand international legal machinery (i.e., multi-lateral and bilateral agreements and conventions) to cover such cooperation. This admittedly is a difficult political struggle, as demonstrated by the resistance of many states to join the Budapest Convention on Cybercrime, which provides for such cooperation. Yet, policy makers will increasingly need to weigh whether some sacrifice of sovereignty and outsiders’ scrutiny of their cyberspace are bearable costs for providing means to resolve cyber conflicts and redress grievances, other than escalation-prone cyber operations.

Furthermore, some existing international institutions with powers to sanction violations of their rules might be appropriate venues for states to respond to exploits and attacks. Most notably, the United States might argue at the World Trade Organization that China’s thefts of American corporations' intellectual property and trade should be punished as violations of WTO rules. International lawyers tend to believe that lacking indisputable evidence, this would be a fruitless step, and, in any case, the WTO and other international regimes are better able to deal with specific incidents rather than the widespread non-compliance issues like those depicted in the allegations.
against China.

However, James Lewis suggests that the WTO rules and practices could nevertheless be leveraged were the United States to hint that since China does not respect them, it might no longer consider itself bound by them. Lewis defends considering this step, which would seem to threaten collapse of the WTO, on legal and pragmatic grounds. The organization’s rules allow a member to take steps that it deems are needed to protect its essential security interests, and China would likely reduce its espionage since it could expect to suffer more than the United States through an ensuing disruption in international trade. Although he characterizes the threat as “raising the possibility of force majeure,” it is less combative than cyber operations, and it is embedded in an international civil institutional order that is worth preserving, albeit not at all costs.

More generally, recommendations of responses to cyber exploits and attacks that would utilize institutions of international cooperation bring into relief several paradigms or prisms for viewing them: threats to international security that might warrant military actions, crimes that are challenges for international policing, and violations of international economic regimes that might be sanctioned. Each paradigm has its associated terms, so, for instance, labeling industrial espionage “economic warfare” can bias decision making toward military type responses. Alternatively, some rhetorical disarmament (i.e., using terms like “intellectual property theft” and “illegal trade practices”) can suggest under types of responses and reduce tendencies to escalate.

States should enlist respected third parties in verifying attributions of exploits and attacks and in mediation of cyber conflicts. Many states feel increasingly threatened cyberspace and are acting more aggressively there, but they lack experience in managing the conflicts that these postures facilitate. The interventions of third parties can therefore help in keeping parties apart and rectifying misunderstandings or misperceptions. Because attributions made by private sector and academic cyber investigators, like Mandiant, Kaspersky Lab, and Citizen Lab, will command greater credibility from the global public than governments’ allegations and denials, their ”naming and shaming” bad actors can have better chances of stopping the behavior. The possibility of referring some cases to such organizations may reduce the feeling among government decision makers of the need to take matters into their own hands. Similarly communication among national CERTs can provide a behind the scenes alternative to confrontation between states for getting explanations and restraints on perceived attacks. Fortunately, states through their representatives at various international forums have already recognized that such third parties have a role in providing security and order in international cyberspace.

**Conclusion.** In discussing the possibilities for cyber norms to promote international security, the former chair of the UN Group of Governmental Experts (UNGGE) wondered what principle of self-constraint would lead states to accept such norms. As noted,
a fear of the consequences of cyber conflicts might not be sufficiently powerful for them to voluntarily sacrifice some claims of sovereignty and desires for freedom of action. Their motivation might be bolstered by their recognizing that in addition to protecting their national assets from cyber exploits

and attacks, they have a duty to assure their citizens’ productive and beneficial uses of cyberspace. In the meantime, practices, which this essay describes, in making choices that avoid conflict and escalation more obvious and less risky can contribute to preserving international security from cyber threats.

NOTES


2 Most notably, the Organization for Security and Cooperation in Europe (OSCE) failed in December 2012, to agree on a previously formulated, weak set of such measures. This was due to a Russian demand for the recognition that certain information transmitted from one state could threaten a second state’s national security on the basis of ideational content, as distinct from the transmission of code (e.g., malware, which can disrupt or damage the second state’s digital networks or appliances dependent on them). Russia’s demand is subsumed by the term “information security” which protects against information operations, including propaganda (or dissidents’ messaging and organizing), whereas the Western term “cybersecurity” refers to protecting operation of the networks.

3 A copy of the text can be found at: http://www.guardian.co.uk/world/interactive/2013/jun/07 obama-cyber-directive-full-text

4 In the United States, such caution was expressed in the Powell Doctrine and exercised in decision-making for the First Gulf War (1991-1992), but not exercised in the Second or Iraq War (2003).


7 According to a 2011 report, the Internet accounts for nearly 4% of GDP in advanced industrial and post-industrial countries and 21% of GDP growth there, see McKinsey Global Institute, Internet matters: the net’s sweeping impact on growth, jobs and prosperity, May, 2011. http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_great_transformer


9 For discussion of the drawbacks in these moves, see Martin Libicki, Crisis and Escalation in Cyberspace, (Santa Monica, CA: RAND Corp., 2012): 33–37. Can be found at: http://www.rand.org/pubs/monographs/MG1215.html


11 China’s position regarding demilitarization has not daunted its efforts to integrate use of cyber attacks on adversaries’ military networks into their strategic military planning. See, for example, Timothy Thomas, Dragon Bytes: Chinese Information–War Theory and Practice from 1995–2003, (Fort Leavenworth, KS: Foreign Military Studies Office, 2004).

12 China’s agreement on this issue is suggested by the U.S. State Department’s June 7, 2013, “Statement on Consensus Achieved by the UN Group of Governmental Experts On Cyber Issues” that praises the group, including a representative of China, for affirming that international law applies to cyberspace. http://www.state.gov/r/pa/prs/ps/2013/06/210418.htm. However, other sources have reported that China is more reticent and has, at most, affirmed the applicability of LOAC, as part of international law, only in bilateral talks with the United States.

14 Whether the cyber operations popularly called Stuxnet, after the worm that enabled penetration of the controls of the centrifuges themselves constituted use of force has been a subject of some debate in the United States, but of much less doubt elsewhere.


16 Arguably a state would gain little by inflating or deflating the severity it assigns to a cyber operation in order to discourage another’s use of that operation or trivialize its own use. By grossly overstating its perception of the threat, the state creates a bluff that tempts other states to call. By grossly understating the threat, it risks other states dismissing its view as non-credible and self-serving.


20 Dima Adamsky, The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US and Israel. (Stanford, CA: Stanford Security Studies, 2010) provides an excellent introduction to these concepts, but does not discuss the role of cyber operations per se. See also
