The Nature of Strategic Instability in Cyberspace

Martin Libicki
Senior Management Scientist
RAND Corporation

The quest for strategic stability was a leitmotif of Cold War thinking. Worry over whether the Soviet Union could disable U.S. nuclear deterrence with a first strike shaped U.S. nuclear posture in the late 1950s, and gave impetus to concerns over a "window of vulnerability" in the late 1970s. Arguments for and against the Anti-Ballistic Missile (ABM) Treaty and strategic arms limitations often cited their putative effects on strategic stability in the early 1970s and throughout the 1980s. Fears of unstable reactions pervaded planning for nuclear command and control as well as indication and warning systems.

Although nuclear arms are still with us, issues associated with strategic stability have spilled over from the nuclear arena to the new arena of cyberspace. Cyberwar—a campaign of non-kinetic attacks on information systems to disrupt and corrupt their operations—is said to present similar stability problems. Although cyberdefenses do exist, unlike nuclear defense, offense in cyberspace is cheap; defenses are expensive; and attackers are constantly getting through. In the first three months of 2011 alone, news reports were punctuated with multiple network penetrations: Australia's government, France's government, large oil companies, the NASDAQ, Sony's PlayStation network, and even RSA, a major security firm. Elite hackers claim that, given the time, they can penetrate any system and wreak mischief. Not for nothing did last July's cover story in *The Economist* picture cyberwar as the digital equivalent of the nuclear bomb, a threat to civilization that necessitated international negotiations and arms control.[1]

Martin Libicki is a senior management scientist at the RAND Corporation, where he focuses on the impact of information technology on security. Prior to joining RAND in 1998, he worked at the National Defense University. He also teaches at Georgetown University and the U.S. Naval Academy.

Martin Libicki

Does cyberwar threaten strategic stability? Do the systemic characteristics of confrontation in cyberspace reward aggression, require responses in which speed takes priority over deliberation, make it difficult for states to back down, and thereby tend to push states into conflicts that neither one sought? In an unstable environment, might chance factors cause some states to react in ways that bring out responses from other states that reinforce mutual fear? Does cyberspace recreate the Prisoner's Dilemma (an unstable two-handed dilemma in which each prisoner rats on the other thereby making both worse off)? Is there, within cyberspace, the equivalent of the railroads whose schedules hastened the alacrity with which both sides mobilized prior to the onset of World War I, lest they find themselves outgunned on the front at the outset?

To date, cyberwar has yet to claim its first life.

I argue no—with one possible exception, noted last.

Nuclear Weapons Trump Cyberspace

The continued existence of nuclear weapons limits the ability of cyberspace to cause strategic destabilization. Although nuclear powers may choose to yield to the will of another state, they cannot be annihilated or taken over, since the cost would exceed anything that cyberwar can wreak. Even forced regime change may be off the table. To date, cyberwar has yet to claim its first life. In other words, a state with nuclear weapons that is worried largely about the survival of the nation and its citizens can afford to ignore whatever relative superiority its rivals may enjoy in cyberspace. Thus, the stakes of cyberwar, although potentially large, are less than existential.

No State can Disarm another State's Cybercapabilities Through Cyberwar Alone

The prospect that a state that strikes first can create an overwhelming presumption of ultimate victory, if true, would make for a highly unstable system, in that anything or nothing at all could lead states to strike first lest they find themselves at the mercy of those who did. The nuclear era was pervaded by the fear that one side or another would gain a decisive first strike capability. The victimized state would either be completely disarmed, or so stripped of strike power that it could not credibly strike back without risking the destruction of its cities to no useful strategic end. The discovery in the mid-1950s that the strategic posture of the United States lent itself to being disarmed by a first strike profoundly

affected nuclear planning thereafter.[2]

By contrast, cyberwar cannot disarm cyberwarriors. The reason is simple. The ability to carry out cyberwar requires no more than four inputs: clever hackers, intelligence on the target's operations and vulnerabilities, a computing device, and network connections. Cyberwar, clearly, cannot destroy the first two, leaving as theoretical possibilities attacks on the attacking computer and the accommodating network. Cyberattacks against an unprepared *attacking* computer (almost an oxymoron) sitting on the network at the time—two less-than-universal conditions—may disable it temporarily, but will not destroy it. Such computers can be rebooted, or, at worst, returned to factory conditions within hours; new ones cost a few hundred dollars at most. Attacking the accommodating network—for instance, by trying to sever the attacking country from the rest of the world electronically— is a theoretical possibility. However, any serious cyberwar-capable state has probably figured out how to carry out cyberattacks starting from someone else's territory or, better yet, from within the territory of the country that it is targeting.[3] It would not be unreasonable to assume that many states capable of waging cyberwar have implanted malware in computers all around the world for the sake of convenience, not to mention survivability. Then, if a crisis starts they would only have to signal the target system—in any one of many ways—to turn the malware on and thereby trigger faults in the targeted computers.

If a first strike cannot disarm its intended target, there is no advantage to going first: when the cyberdust clears, the attacker is no better off than before, even relative to its foe. Indeed, it may be worse off a week later. The best cyberattackers come equipped with knowledge of vulnerabilities in their target computers. Typically, such vulnerabilities are generic: for example, they affect all computers running a particular class of software. Those who attempt a first strike usually create enough evidence in the target computer to allow a computer owner to figure out what kind of vulnerabilities was exploited. Such vulnerabilities can then be patched—leaving the attacker with that many fewer tricks in its arsenal. [4]

A variant on the strategy of carrying out a disarming first strike and thereby acting without fear of reprisal might be to carry out a first strike and then isolate oneself from the rest of the world to ward off reprisals. Richard Clarke and Robert Knake posited such a scenario in their book *Cyber War*.[5] This seems akin to taking a quick punch at someone and then jumping into the water to avoid a counterpunch: not only does the aggressor get quite wet, but it has to emerge eventually—at which point it is no less vulnerable. When the likelihood

that the effects of most cyberattacks are temporary is taken into account, the attractiveness of this strategy further dwindles. There is nothing in this ploy that would convince potential victims of such an attack to preempt the attacker by going first. More likely, if potential target states are able to anticipate as much, they could engineer a possible response that would work even if the attacking country managed to disconnect itself from the Internet.

Cyberwar against Nuclear Capabilities Should not be Capable of Creating Instability

Although cyberattacks could create serious instability by altering the terms of a nuclear standoff, it is hard to believe that protecting its networks against someone exploiting their vulnerabilities could rise beyond the level of yet one more thing for states to keep in mind when designing a nuclear establishment. Operations in cyberspace are rarely capable of breaking computer systems or, for the most part, devices that such systems control. True, the Stuxnet worm did apparently shorten the lifespan of Iran's nuclear centrifuges, but it targeted machinery actively subject to computer commands—which were then interfered with (after all the machines in question were in use at the time). Nuclear weapons, by contrast, are less vulnerable to remote sabotage.

Conceivably, one state could hack into the nuclear command system of another state, render its weapons unreachable, and use the temporary monopoly of power to coerce its target. But states are extraordinarily cautious in the construction of their nuclear systems and usually give primary command and control a great deal of thought and redundancy.[6] Even technologically unsophisticated states retain very simple but robust ways of wielding their nuclear weapons if primary systems fail. Furthermore, it is unclear how the aggressor state would know that its cyberoperations had, in fact, disabled the target state's ability to fire its nuclear weapons. Its hackers may have disconnected everything they found, but how confident would they be that they found everything? There could be several duplicated paths, each separate from the other. Command-and-control software, unlike missile silos or submarine pens, is invisible from space.

Finally, in the event of a successful cyberattack, the target state's nuclear weapons are not destroyed, only disconnected for a time. Unless the aggressor can quickly disarm the target state's nuclear capacity on its own (in which case, the nuclear standoff alone is unstable, with or without cyberwar), the target state need only bargain for enough time to reconnect its weapons to get back into the deterrence game.[7]

Cyberwar does not Lend itself to Mutually Reinforcing Alert–Reaction Cycles

During times of heightened Cold War tension, warnings of possible nuclear activities by one side would cause the other side to raise its alert level.[8] Raising the alert level led to procedures that would make nuclear weapons more readily usable (e.g., aircrews would be recalled to base, submarines would head to sea). The other side would perceive that its rival was moving closer to striking first and would raise *its* alert level, prompting the first side to move even closer to striking. Such a dynamic reflects the world of nuclear warfare in which the only response to a growing threat was to raise the readiness of one's offense. Second, there was a decided first-strike advantage. Third, many reactions were visible.

Unlike nuclear war, raising cyberdefenses rather than offenses is quite a viable reaction to the heightened threat of cyberwar.

By contrast, the first-strike advantage in cyberspace is minimal. Furthermore, a great deal of what goes on in cyberspace is not readily visible. Being invisible, such operations are less likely to cause rapid reactions. Unlike nuclear war, raising cyberdefenses rather than offenses is quite a viable reaction to the heightened threat of cyberwar. Such defenses may include selectively disconnecting systems, disallowing certain services, tightening access controls, or heavily filtering what enters and leaves networks. Although greater defenses might presage a turn to hostilities, the linkage is relatively weak and could be explained by a third-party cyberthreat, the discovery of a novel threat vector, or simply the belated realization that cyberdefenses need to receive proper attention.[9]

Cyberdefenses are not Inherently Destabilizing

The relatively distinct nature of defensive and offensive measures in cyberspace also casts doubt on the destabilizing effects of elevating defenses. Many of the objections to building ballistic missile defenses assumed any country that had rendered other countries' missiles "impotent and obsolete" would be able to attack other nations with impunity and, having achieved that position, would be able to lash out whenever it wished.[10] Analogously, it could be argued that a nation that had perfected the art of cyberdefenses could hold other countries hostage to its offensive cyberwar capability. This proposition also holds little water, even if one ignores the fact that nuclear weapons and advanced conventional weapons can trump cyberweapons.

First, good defenses are unlikely to be exclusive to one country. Instead, some will proliferate from kindness, freely distributed by people who assume that frustrating cyberattacks is always a good thing just as frustrating crime is always a good thing. Many of the institutions that exist within the one country also have branches (and support contractors) overseas; improvements for them at home would inevitably be improvements for them (and their support contractors) overseas. Moreover, hackers who chance upon these defenses and learn about innovations in defenses, if not necessarily how they work, are also likely to spread knowledge. Second, confidence in such defenses may be hard to come by. Cyberwar is one surprise after another—inevitably so, because it depends on the exploitation of vulnerabilities that the hacker has found and that the defender has not (if the defender knew about its own vulnerabilities, it would undoubtedly fix them). Confidence in defenses requires confidence that one will never be surprised again. Third, a state cannot wield power over other states until it can convince them that its own confidence is merited. Rivals could easily conclude that no such confidence is possible in cyberspace. Fourth, even if the foe is sufficiently fearful to yield to coercion, it may also be prepared to make many tradeoffs to bolster its own defenses (for instance, by spending money and creating inconvenience to users). It is willing to do this at least to the point where the worst possible damage it can bear from a cyberattack is below the threshold at which it feels it must yield to the aggressor state. All told, there is no rational reason to become alarmed because a potential adversary's defenses are getting better.

Arms Races are Hard to Conduct in the Dark

Might instability arise from an arms race, particularly if it is large enough to bankrupt the participants? In a sense, such a race already exists, not between states but between defense and offense. As old vulnerabilities are fixed, new ones are discovered either in old software, new software, or in new uses for software that violate implicit assumptions about security (e.g., opening up an e-mail could not introduce malware into one's computer). As anti-virus firms scour the web to identify more malware signatures, malware producers have developed morphing technologies that constantly create hitherto unseen signatures. Such contests would continue even if states had no interest in cyberwar—although the fact that states are interested means that additional resources are being poured into both sides of that contest.

Nevertheless, the notion that states have to develop offensive cyberweapons

simply because their rivals do so has little basis in theory or fact. First, states have little knowledge of exactly what weapons are in the arsenal of their rivals. Indeed, if they actually knew what weapons their foes had, they might well know the vulnerabilities of their own weapons and fix such vulnerabilities. Second, the best response to an offensive weapon is a defensive weapon, not another offensive weapon. A defensive weapon can neutralize the attack, but an offensive weapon cannot except under unusual circumstances. Third, the whole notion of an offense-versus-offense dynamic requires that the underlying attack and retaliation actually make sense as a war-fighting and war-wining strategy. Were that so, deterrence would be primary, but deterrence is a very difficult notion in cyberspace because states have a variety of ways to hide their cyberattack tracks and more than enough motive to do so.[11] Fears of an arms race simply have no good basis in logic. Incidentally, even if they did, it is hard to imagine how an arms race in cyberspace could come close to having a major economic impact. The intellectual skills required to compete in this contest are so specialized that states will run out of such people long before they run out of money to pay them.

It is hard to imagine how an arms race in cyberspace could come close to having a major economic impact.

Arms Control does not Contribute Significantly to Strategic Stability

Banning cyberweapons requires stretching the concept of *weapons* beyond its bursting point. A cyberattack requires two components: knowledge of the target and its vulnerabilities, and a capability to translate such knowledge into malware or other attack methods that succeed in getting the system to execute the attacker's commands while evading detection. Generally, the first is harder. An attacker needs to know one or more zero-day vulnerabilities—problems in software of which their producers are unaware—against a well-run system. Yet, knowledge per se is not a weapon in the sense that we think of weapons. Weaponization is the simpler half, and there is a great deal of material in existence that can be used to develop offensive capabilities once vulnerabilities are found.[12] Banning cyberweapons is akin to banning a state's ability to build fertilizer bombs: fertilizer is legitimate to own, and the cases and fuses, while easier to proscribe, are also easier to acquire once the time comes to make a bomb. In the end, the controllers have very little leverage.

Additionally, cyberarms control does not provide the stability that nuclear arms control did during the Cold War. If a state is known to lack nuclear weapons

or a nuclear weapons program, its rivals can assure themselves that it would take years before they would have to worry about facing such weapons in a crisis. The same cannot be said of cyberweapons, since nearly everything about offensive cyberwar is quite well hidden, and the time required to go from no capability to a respectable one in cyberwar is relatively short.

Cyberwar May Contribute to Strategic Instability in Scenarios where Getting a Jump can be Decisive

A more plausible concern may be the effect of cyberwar on non-nuclear scenarios. Consider two countries: one is attempting to take an offshore island, while the other hopes to get to the island fast enough and with sufficient forces to rescue it from aggression. The potential invader hopes to delay its rival's arrival by a few days, and a successful cyberattack might well retard the other side's progress into the theater. If so, the aggressor can finish the takeover before it has to confront its rival.[13] In the hopes that the rival's forces show up too late, it carries out a no-warning cyberattack on the rival's logistics or mobilization system prior to a no-warning invasion.

The "no-warning" part of the equation speaks to the way a capability for cyberwar can exacerbate a potential instability. Although this scenario is a plausible one, it is also a dicey ploy by the invader, who must have great confidence in the efficacy of cyberwar before launching an invasion. Finally, although one real-world context for such a scenario can be imagined, coming up with a second is not so easy.

Conclusion

Although cyberwar is not necessarily good for civilization, fears of its destabilizing effect are exaggerated. Besides being easily trumped by nuclear weapons and most forms of conventional conflict, the primary characteristics of cyberwar—its temporary effects, covert characteristics, the impossibility of disarming cyberwarriors, and the usefulness and primacy of cyberdefense—do not lend themselves to rapid uncontrollable action–reaction cycles. Cyberwar may take place at the speed of light, but this hardly suggests that strategic decisions about the use of cyberwar have to take place any faster than the speed of understanding.

Notes

1. See the cover of its July 2, 2010 issue.

2. Albert Wholstetter, "The Delicate Balance of Terror," *Foreign Affairs* 37, No. 2 (1959): 211-23. This article was based in part on an analysis of U.S. bomber basing in the mid-1950s.

3. Chinese military writings call for cyberattacks against the United States to be launched from within the United States. See: James Mulvenon, "Information Warfare and China's Cyber-warfare Capabilities," (Speech, Carnegie Endowment for International Peace, Washington, DC, February 10, 2011).

4. Microsoft learned that the Stuxnet virus used four zero-day vulnerabilities in July, 2010 (roughly two months before the news hit the *New York Times*). It patched the most critical one (the LNK vulnerability) within twenty days by August 2. See: Ryan Naraine, "As Attacks Escalate, Microsoft Ships Emergency Windows Patch," *Zero Day Blog*, August 2, 2010, www.zdnet.com/blog/security/as-attacks-escalate-microsoft-ships-emergency-windows-patch/7027. The second patch, associated with a bug on the Print Spooler Service, was patched on September 14. See: Ryan Naraine, "Stuxnet Attackers used 4 Windows Zero-Day Exploits," *Zero Day Blog*, September 14, 2010, www.zdnet.com/blog/security/stuxnet-attackers-used-4-windows-zero-day-exploits/7347. The last two were patched on December 10. See: Lucian Constantin, "Microsoft to Patch IE and Stuxnet 0-Day Vulnerabilities Next Tuesday," *Softpedia News*, news.softpedia.com/news/Microsoft-to-Patch-IE-and-Stuxnet-0-Day-Vulnerabilities-Next-Tuesday-171996.shtml).

5. Richard Clarke and Robert Knake, *Cyber War: The Next Threat to National Security and What to Do About It*, (New York: Harper Collins, 2010), 179-218.

6. The 1983 movie *War Games*, besides offering a cautionary tale against nuclear war in general and automation of military functions in particular, also suggested that connecting nuclear command systems to a modem bank is inadvisable.

7. One can imagine hypothetical circumstances under which one state would target another state's nuclear weapons and then use cyberattacks to immobilize a retaliatory response while the nuclear weapons that were not destroyed in the first round are sought and destroyed over the next few days. That presumes 1) that the aggressor has enough weapons to silence the target's retaliatory capability sufficiently, and 2) that the remaining weapons can be found in the relatively short time afforded by the immobility subsequent to a cyberattack. Neither assumption would apply to the U.S.-U.S.S.R standoff, where neither side had enough weapons for assured complete destruction of the other, and a large share of the weapons that were not destroyed were in submarines and very difficult to find.

8. This argument is made in Paul Bracken, "Strategic War Termination," in *Managing Nuclear Operations,* ed. Ashton B. Carter et al. (Washington, DC: The Brookings Institution, 1987).

9. In the Cold War, the third party nuclear threat to the United States was far less consequential than the Soviet nuclear threat. Today's cyberwar environment features three comparably competent states, more strong second-tier states, and a serious transnational criminal capability.

10. From President Reagan's March 23, 1983 speech announcing the Strategic Defense Initiative. See: Ronald Reagan, "Address to the Nation on Defense and National Security," (Speech, Washington DC, March 23, 1983), http://www.reagan.utexas.edu/archives/speeches/1983/32383d.htm.

11. Martin Libicki, "Why Cyberdeterrence is Different," in *Cyberdeterrence and Cyberwar*, (Santa Monica, CA: RAND Corporation, 2009): 39-74.

12. Analogy may be drawn to nuclear weapons programs. No state that has amassed the requisite fissile material has failed to complete all the other weaponization steps on the way to building a bomb. See: Peter D. Zimmerman, "Proliferation: Bronze Medal Technology is Enough," *Orbis* 38, no. 1 (1994): 67.

13. Cleaning out the original malware from every single machine is a much longer process, but if the malware cannot be signaled from the outside or if its effects can be counteracted in other ways (e.g., through more human oversight), the system owners can (warily) use the system in the interim.