Offensive Cyber Operations and the Use of Force

Herbert S. Lin[*]

INTRODUCTION

Hostile actions against a computer system or network can take two forms.[1] One form – a cyber attack – is destructive in nature. An example of such a hostile action is erasure by a computer virus resident on the hard disk of any infected computer. In this article, "cyber attack" refers to the use of deliberate actions and operations – perhaps over an extended period of time – to alter, disrupt, deceive, degrade, or destroy adversary computer systems or networks or the information and (or) programs resident in or transiting these systems or networks.[2] 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 the adversary's computer systems and networks to be unavailable or untrustworthy and therefore less useful to the adversary.

The second form – cyberexploitation – is nondestructive. An example is a computer virus that searches the hard disk of any infected computer and emails to the hostile party all files containing a credit card number. "Cyberexploitation" refers to the use of actions and operations – perhaps over an extended period of time – to obtain information that would otherwise be kept confidential and is resident on or transiting through an adversary's computer systems or networks. Cyberexploitations are usually clandestine and conducted with the smallest possible intervention that still allows extraction of the information sought.[3] They do not seek to disturb

[*] Chief Scientist, Computer Science and Telecommunications Board, National Research Council (NRC) of the National Academies. At the NRC, Dr. Herbert Lin has been study director of major projects on public policy and information technology. Prior to his NRC service, he was a staff member and scientist for the House Armed Services Committee (1986-1990), where his portfolio included defense policy and arms control issues.

[1] This article is based almost entirely on material drawn from NATIONAL RESEARCH COUNCIL, TECHNOLOGY, POLICY, LAW, AND ETHICS REGARDING U.S. ACQUISITION AND USE OF CYBERATTACK CAPABILITIES (William A. Owens, Kenneth W. Dam & Herbert S. Lin eds., 2009) [hereinafter NRC Report]. The project was supported by the MacArthur Foundation and the Microsoft Corporation, although the views reflected in this article and in the NRC Report do not necessarily reflect the views of either of these sponsors. The NRC Report covers a host of issues concerning the Law of Armed Conflict (LOAC) that are not discussed in this article, most notably how cyber attack might be treated under *jus in bello*.

[2] 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.

[3] If the requirement for stealth is met, the adversary is less likely to take countermeasures to negate the loss of the exfiltrated information. In addition, stealthiness enables penetration of an adversary's computer or network to result in multiple exfiltrations of intelligence information over the course of the entire operation.

the normal functioning of a computer system or network from the user's point of view, and the best cyberexploitation is one that a user never notices.

For purposes of this article, the term "offensive cyber operations" will include military operations and activities in cyberspace for cyber attack against and (or) cyberexploitation of adversary information systems and networks. When greater specificity is needed, the terms "cyber attack" and "cyberexploitation" will be used.[4]

Although the objectives and the legal and policy constructs relevant to cyber attack and cyberexploitation are quite different (see the table in the Appendix to this article), the technological underpinnings and associated operational considerations of both are quite similar.

Cyber attacks and cyberexploitation require a vulnerability, access to that vulnerability, and a payload to be executed.[5] In a noncyber context, a vulnerability might be a lock to a file cabinet that could be easily picked. Access would be an available path for reaching the file cabinet. From an intruder's perspective, access to a file cabinet located on the International Space Station would pose a very different problem from that posed by the same cabinet located in an office in Washington, D.C. The payload is responsible for executing the action taken by the intruder after the lock is picked. For example, the intruder can destroy the papers inside, or alter some of the information in those papers.

The primary technical difference between cyber attack and cyberexploitation is in the nature of the payload to be executed – a cyber attack payload is destructive whereas a cyberexploitation payload acquires information nondestructively. In addition, because a cyberexploitation should not be detected, the cyber operation involved must only minimally disturb the normal operating state of the computer involved. In other words, the intelligence collectors need to be able to maintain a clandestine presence on the adversary computer or network despite the fact that information exfiltrations provide the adversary with opportunities to discover that presence.

[4] Although there is no published Department of Defense (DoD) definition for "offensive cyber operations," the U.S. Air Force uses the term to encompass attack and exploitation of adversary information systems. *See* Air Force Research, Development, Test and Evaluation (RDT&E) Budget Item Justification for FY 2010, Appropriation/Budget Activity 3600 – Advanced Technology Development (ATD) 1 (May 2009), *available at* http://www.dtic.mil/descriptivesum/Y2010/AirForce/0603789F.pdf ("The Battlespace Information Exchange project will . . . demonstrate offensive cyber operations technologies allowing attack and exploitation of adversary information systems by the Air Force.").

[5] In the lexicon of cybersecurity, "using" or "taking advantage" of a vulnerability is often called "exploiting a vulnerability." The term "cyberexploitation" in an espionage context is a cyber offensive action conducted for the purpose of obtaining information. The context of usage will usually make clear which of these meanings of "exploit" is intended.

I. TECHNOLOGY OF OFFENSIVE CYBER OPERATIONS

A. Vulnerabilities

For a computer or network, a vulnerability is an aspect of the system that can be used to compromise that system (for illustrative vulnerabilities, see the Appendix). "Compromise" is used here as a verb meaning to attack or exploit. Weaknesses may be introduced accidentally through design or implementation flaws. A defect or "bug" may open the door for opportunistic use of that vulnerability by an adversary. Many vulnerabilities are widely publicized after discovery and may be used by anyone with moderate technical skills until a patch can be disseminated and installed.[6] Adversaries with the time and resources may also discover unintentional defects that they protect as valuable secrets, also known as zero-day exploits.[7] As long as those defects go unaddressed, the vulnerabilities they create may be used by adversaries.

Vulnerabilities may also be introduced intentionally. Of course, vulnerabilities are of no use to an adversary unless the adversary knows they are present on the system or on the network being compromised. But an adversary may have some special way of finding vulnerabilities, and nation states in particular often have special advantages in doing so. For example, although proprietary software producers jealously protect their source codes as intellectual property upon which their businesses are dependent, some such producers are known to provide source code access to governments under certain conditions.[8] Availability of source code for inspection increases the likelihood that the inspecting party will be able to identify vulnerabilities not known to the general public. Furthermore, through covert and nonpublic channels, nation states may even be able to

[6] The time lag between dissemination of a security fix to the public and its installation on a specific computer system may be considerable, and is not always due to unawareness on the part of the system administrator. Sometimes 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.

[7] 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. These types of remotely accessible zero-day attacks on services appear to be less frequently found as time goes on. In response, a shift in focus to the client side has occurred, resulting in many recent zero-day attacks on client-side applications. *See* DANIEL GEER, MEASURING SECURITY 1, 278-287 (2006), *available at* http://www.geer.tinho.net/measuringsecurity.tutorialv2.pdf (providing data and analyses of zero-day attack trends).

[8] *See, e.g.*, Microsoft.com, Government Security Program, http://www.microsoft.com/industry/publicsector/government/programs/GSP.mspx.

persuade vendors or willing employees of those vendors to insert vulnerabilities – secret "back doors" – into commercially available products (or require such insertion as a condition of export approval), by appealing to their patriotism or ideology, by bribing, blackmailing, or extorting them, or by applying political pressure.

B. Access

In order to take advantage of a vulnerability, an adversary must have access to it. Targets that are "easy" to compromise are those that involve relatively little preparation on the part of the adversary and where access to the target can be gained without much difficulty, such as a target that is known to be connected to the Internet. "Difficult" targets require a great deal of preparation on the part of the adversary, and access to the target can be gained only with great effort, or may even be impossible for all practical purposes. For example, the onboard avionics of an adversary's fighter plane are not likely to be connected to the Internet for the foreseeable future, which means that launching a cyber attack against it will require some kind of close access to introduce a vulnerability that can be used later. In general, it would be expected that an adversary's important and sensitive computer systems or networks would fall into the category of difficult targets.

Access paths to a target may be intermittent. For example, a submarine's on-board administrative local area network would necessarily be disconnected from the Internet while underwater at sea but might be connected to the Internet while in port. If the administrative network is ever connected at sea to the on-board operational network, which controls weapons and propulsion, a useful though intermittent access path may be present for an adversary.

Access paths to a target can suggest a way of differentiating between two categories of compromise:

- **Remote access**: Where a compromise is launched at some distance from the adversary computer or network of interest. The canonical example of a remote access compromise is using the access path provided by the Internet, but other examples might include accessing an adversary computer through a dial-up modem attached to it or through penetration of the wireless network to which it is connected.

- **Close access**: Where a compromise takes place through the local installation of hardware or software functionality by friendly parties (e.g., covert agents, vendors) in close proximity to the computer or network of interest. Close access is a possibility anywhere in the supply chain of a system that will be deployed. It may well be easier to gain access to the system before it is deployed.

C. Payload

"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, and destroy or alter files on the system.

Payloads can have multiple capabilities when inserted into an adversary system or network; they can be programmed to do more than one thing. The timing of these actions can also be varied, and if a communications channel to the adversary is available, payloads may be remotely updated. Indeed, in some cases, the initially delivered payload consists of nothing more than a mechanism for scanning the system to determine its technical characteristics and another mechanism through which the adversary can deliver the best software updates to further the compromise.

D. Effects

Cyberexploitations target the confidentiality of information stored on or passing through a system or a network. Under normal circumstances, such information should be available only to authorized parties. A successful cyberexploitation compromises the confidentiality of such information and makes the information available to the adversary.

Cyber attacks (as opposed to cyberexploitations) target one of several attributes of these components or devices and seek to cause a loss of integrity, a loss of authenticity, or a loss of availability, which includes theft of services:

- **Integrity**: A compromise of integrity refers to the alteration of information (a computer program, data, or both) so that under some circumstances of operation, the computer system does not provide the accurate results or information that one would normally expect even though the system may continue to operate.

- **Authenticity**: A compromise of authenticity obscures or forges the source of a given piece of information. A message whose authenticity has been compromised will fool a recipient into thinking it was properly sent by the asserted originator.

- **Availability**: A compromise in availability means that the functionality provided by the target system or network is not available to the user: email sent by the targeted user does not go through, the target user's computer simply freezes, or the response time for that computer becomes intolerably long, possibly leading to catastrophe if a physical process is being controlled by the system.

The compromises above are direct effects of a cyber attack. In addition, cyber attacks may result in indirect effects on the systems and (or) devices that the attacked computer system or network controls or interacts with, or on the people who use or rely on the attacked computer system or network. For example, an adversary's electric power grid may be controlled by computer. An attack on the grid's computers may have effects on the power grid itself – indeed, producing those indirect effects on the grid may be the primary purpose of the attack. Furthermore, because virtually anything can be connected to a computer system or network, the scope and nature of effects resulting from a cyber attack can span an enormous range. The indirect effects of a cyber attack are almost always more important to the attacker than the direct effects, although both direct and indirect effects must be taken into account when ascertaining the significance of a cyber attack.

II. POSSIBLE OBJECTIVES FOR OFFENSIVE CYBER OPERATIONS

What might cyberexploitations seek to accomplish? The following paragraphs describe hypothetical examples of cyberexploitation:

- **Exploit information available on a network**. For example, a cyber operator might monitor passing network traffic for keywords such as "nuclear" or "plutonium," and copy and forward to the cyber operator's intelligence services any messages containing the words for further analysis. A cyberexploitation against a military network might seek to exfiltrate confidential data indicating orders of battle, operational plans, and so on. Alternatively, passwords are often sent in unencrypted form through email, and those passwords can be used to penetrate other systems. This objective is essentially the same as that for all signals intelligence activities – to obtain intelligence information on an adversary's intentions and capabilities.

- **Be a passive observer of a network's topology and traffic**. Networks can be passively monitored to identify active hosts as well as to determine the operating system and/or service versions through signatures in protocol headers, the way sequence numbers are generated, and so on.[9] The cyber operator can map the network and make inferences about important and less important nodes on it simply by performing traffic analysis to determine what the organizational structure is and who holds positions of authority. Such information may be subsequently used to disrupt its operational functionality. If the cyber operator is able to read the contents of traffic (which is likely if the adversary believes the network is secure and thus has not gone to the trouble of encrypting traffic), he can gain much more information about matters of significance to the network's operators. A map of the network is just as important to provide useful information for a cyber attacker, who can use this information to perform a more precise targeting of later attacks on hosts on the local network, which are typically behind firewalls and intrusion detection and prevention systems that might trigger alarms.

- **Conduct industrial espionage**. For example, two former directors of the Direction Générale de la Sécurité Extérieure (DGSE), the French intelligence service, have publicly stated that one of the DGSE's top priorities is to collect economic intelligence. In a September 1991 NBC news program, Pierre Marion, former DGSE director, revealed that he had initiated an espionage program against U.S. businesses for the purpose of keeping France internationally competitive.[10]

What might a cyber attacker seek to accomplish? The following are some examples of potential cyber attacks:

- **Destroy data on a network or a system connected to the network**. For example, a cyber attacker might seek to delete and permanently erase all data files or to reformat and wipe clean all hard disks that it can find. Moreover, destruction of a network also has negative consequences for anything connected to it. For example, power generation facilities controlled by a network are likely to be adversely affected by a disabled network.

- **Be an active member of a network and generate bogus traffic**. For example, a cyber attacker might wish to

[9] *See* Annie De Montigny-Leboeuf & Frederic Massicotte, *Passive Network Discovery for Real Time Situation Awareness* (Apr. 2004), *available at* http://ftp.rta.nato.int/public//PubFullText/RTO/MP/RTO-MP-IST-041///MP-IST-041-14.pdf.

[10] *See* NATIONAL RESEARCH COUNCIL, CRYPTOGRAPHY'S ROLE IN SECURING THE INFORMATION SOCIETY 33 n.1 (Kenneth W. Dam & Herbert S. Lin eds., 1996).

masquerade as the adversary's national command authority or as another senior official or agency and issue phony orders or pass faked intelligence information. Such an impersonation (even under a made-up identity) might well be successful in a large organization in which people routinely communicate with others that they do not know personally. An impersonation objective can be achieved by a cyber attacker taking over the operation of a trusted machine that belongs to the agency or entity of interest (e.g., the national command authority) or by obtaining the relevant keys that underlie their authentication and encryption mechanisms and setting up a new node on the network that appears to be legitimate because it exhibits knowledge of those keys.

- **Clandestinely alter data in a database stored on the network**. For example, the logistics deployment plan for an adversary's armed forces may be driven by a set of database entries that describe the appropriate arrival sequence of various items such as food, fuel, vehicles, and so on. A planner relying on a corrupted database may well find that deployed forces have too many of certain items and not enough of others. The planner's confidence in the integrity of the database may also be affected.

- **Degrade or deny service on a network**. A cyber attacker might try to degrade the quality of service available to network users by flooding communications channels with large amounts of bogus traffic. A denial-of-service attack on the wireless network (e.g., a jamming attack) used to control a factory's operations might well shut it down. Taking over a telecommunications exchange might give a cyber attacker the ability to overwhelm an adversary's defense ministry with bogus phone calls and make it impossible for its employees to use its telephones to do any work. A denial-of-service attack might be used to prevent an adversary from using a communications system and thereby force him to use a less secure method for communications against which a cyberexploitation could be successful.

III. OFFENSIVE CYBER OPERATIONS, THE U.N. CHARTER, AND THE INTERNATIONAL LAW OF ARMED CONFLICT

Offensive cyber operations are potentially relevant to a broad spectrum of international conditions. At one end, relative peace between two nation state competitors reigns with no shooting conflicts or tensions that signal that such conflicts are imminent. Relations between the nation states in question are at least superficially good, although they still compete for advantage in a variety of ways. At the other end is overt and open armed conflict with kinetic weapons. In between is a broad transitional domain between peace and war characterized by terms such as unfriendliness, tension, sanctions, and crisis.

This article addresses some of the legal considerations affecting the domain of behavior that precedes U.S. decisions to use military force, whether cyber or kinetic. The particularly interesting issue of "active defense," the use of cyberforce to neutralize an adversary's cyber attack, is treated by David Graham in his article in this issue.[11] Active defense is also discussed in the NRC Report.

A. The Basic Framework

The Law of Armed Conflict (LOAC) provides the primary legal framework within which to understand constraints on the use of offensive cyber operations.

The LOAC addresses two separate questions. First, when is it legal for one nation to use force against another? This body of law is known as *jus ad bellum*. Second, what are the rules that govern the behavior of combatants who are engaged in armed conflict? This body of law is known as *jus in bello*. It is separate and distinct from *jus ad bellum*.

Jus ad bellum is governed by the United Nations Charter, interpretations of the Charter, and customary international law developed in connection with and sometimes prior to the Charter. Article 2(4) of the Charter prohibits nations from "the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations."

The U.N. Charter contains two exceptions to this prohibition on the use of force. First, Articles 39 and 42 permit the Security Council to authorize uses of force in response to "any threat to the peace, breach of the peace, or act of aggression" in order "to maintain or restore international peace and security." Second, Article 51 provides: "Nothing in the present Charter shall impair the inherent right of individual or collective self-defense if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security." The self-defense contemplated by Article 51 does not require Security Council authorization.

The U.N. Charter does not formally define "use of force," "threat" of force, or "armed attack." Based largely on historical precedents, nations appear to agree that a variety of unfriendly actions, including unfavorable trade decisions, space-based surveillance, boycotts, severance of diplomatic

[11] *See* David E. Graham, *Cyber Threats and the Law of War*, 4 J. NAT'L SECURITY L. & POL'Y 87 (2010).

relations, denial of communications, espionage, economic competition or sanctions, and economic and political coercion do not rise to the threshold of a use of force, regardless of the scale of their effects. Armed attack is likely to include declared war, occupation of territory, naval blockade, and the use of armed force against territory, military forces, or civilians abroad. However, there are no precedents for how offensive cyber operations should be regarded.

Traditional espionage per se does not appear to violate international law. Hays Parks wrote:

> Each nation endeavors to deny intelligence gathering within its territory through domestic laws. . . . Prosecution under domestic law (or the threat thereof) constitutes a form of denial of information rather than the assertion of a *per se* violation of international law; domestic laws are promulgated in such a way as to deny foreign intelligence collection efforts within a nation's territory without inhibiting that nation's efforts to collect intelligence about other nations. No serious proposal has ever been made within the international community to prohibit intelligence collection as a violation of international law because of the tacit acknowledgement by nations that it is important to all, and practiced by each.[12]

B. Applying Jus ad Bellum to Offensive Cyber Operations

This section addresses some of the issues that might arise in applying international law to offensive cyber operations prior to the outbreak of acknowledged armed conflict – when *jus ad bellum* defines the applicable legal regime.[13] Some issues arise when a nation is the target of one or multiple offensive cyber operations and must consider legal issues in formulating an appropriate and effective response. Other issues arise when a nation wishes to launch a cyber attack against another party prior to the outbreak of hostilities but without intending to give the other side a legal basis for regarding its action as starting a general state of hostilities.[14]

To be fair, many similar issues arise when kinetic weapons are used in conflict. The instruments of offensive cyber operations are newer and have

[12] W. Hays Parks, *The International Law of Intelligence Collection, in* NATIONAL SECURITY LAW 433, 433-434 (John Norton Moore et al. eds., 1990).

[13] For example, during acknowledged armed conflict (notably when kinetic and other means are also being used against the same target nation), cyber attack is governed by all the standard LOAC criteria of *jus in bello* – military necessity, proportionality, distinction, and so on. Nevertheless, because of the novelty of such weapons, there will be uncertainties in how the LOAC will apply in any given instance, and such uncertainties may well complicate operational decisionmaking. *See generally* NRC Report, *supra* note 1, at Chapter 7.

[14] *See* DEPARTMENT OF DEFENSE, OFFICE OF GENERAL COUNSEL, AN ASSESSMENT OF INTERNATIONAL LEGAL ISSUES IN INFORMATION OPERATIONS 5 (2d ed. 1999).

certain characteristics not shared with kinetic weapons. This suggests that fewer precedents and analyses are available and that the application of the LOAC and U.N. Charter principles may not be as straightforward as when kinetic weapons are involved.

Although cyber operations are relatively new, offensive cyber operations are potentially subject to the LOAC and U.N. Charter law. LOAC precepts regarding *jus ad bellum* continue to have validity in a cyber context, though there may be ambiguities in how they should be applied in any given situation. The assumption of this article is that the effects rather than the modality of an action are the appropriate starting point for understanding how *jus ad bellum* and the U.N. Charter apply to offensive cyber operations.

Effects-based analysis suggests that the ambiguities are fewest when cyber attacks cause physical damage to property and loss of life in ways that are comparable to kinetic attacks and traditional war, because traditional LOAC provides various relevant precedents and analogies. The ambiguities multiply in number and complexity when a cyber attack does not cause physical damage or loss of life though there may be other negative effects to the nation.[15]

When does a cyber attack constitute a use of force or an armed attack? As a number of analysts have noted,[16] the relevant question is not so much whether a cyber attack constitutes a use of force but rather whether a cyber attack *with a specified effect* constitutes a use of force. That is, the effects of a given cyber attack are the appropriate point of departure for an analysis of this question rather than the specific mechanism used to achieve these effects.

Therefore, if both the direct and indirect effects to be produced by a cyber attack would, if produced by other means, constitute an armed attack in the sense of Article 51 of the U.N. Charter, it is likely that the cyber attack would be treated as an armed attack. Similarly, if a cyber attack had the same effects and was otherwise similar to government-initiated coercive or harmful actions that are traditionally and generally not treated as the "use of force" (e.g., economic sanctions, espionage, or covert actions such as planting information or influencing elections), such a cyber attack would likely not be regarded as an action justifying a use of force in response. This suggests that the term "cyber attack" should be understood as a

[15] *See* Jason Barkham, *Information Warfare and International Law on the Use of Force*, 34 N.Y.U. J. INT'L L. & POL. 57, 84-85 (2001).

[16] *See* Michael Schmitt, *Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework*, 37 COLUM. J. TRANSNAT'L L. 885 (1999); *see also* Barkham, *supra* note 15; DEPARTMENT OF DEFENSE, *supra* note 14. A 1963 exposition by Ian Brownlie discusses a "results-oriented" approach, but without reference to cyber attacks per se. Ian Brownlie, *International Law and the Use of Force by States*, 22 CAMBRIDGE L.J. 144 (1963).

statement about a methodology for action – and that alone – rather than as a statement about the scale of the action's effect.

Some of the issues raised in applying *jus ad bellum* in the context of offensive cyber operations are discussed below.

1. Criteria for Defining "Use of Force"[17]

The U.N. Charter was formulated in an era that predates the advent of cyber attack and thus its notions of use of force and armed attack do not take into account the dependence of modern society on the existence and proper functioning of an extensive infrastructure that itself is increasingly controlled by information technology. Actions that significantly interfere with the functionality of that infrastructure can reasonably be regarded as uses of force, whether or not these actions cause immediate physical damage. Thus, cyber attacks on the controlling information technology for a nation's infrastructure that has a significant impact on the functioning of that infrastructure (whether or not it caused immediate large-scale death or destruction of property) would be an armed attack for Article 51 purposes, just as would a kinetic attack that managed to shut down the system without such immediate secondary effects.

The following examples illustrate possible scenarios that raise questions about the appropriate definition of a use of force:

- **A cyber attack temporarily disrupts Zendia's stock exchanges and makes trading impossible for a short period**. Bombs dropped on Zendia's stock exchanges at night, so that casualties were minimized, would be regarded as a use of force or an armed attack by most observers, even if physical backup facilities were promptly available so that actual trading was disrupted only for a few hours. The posited cyber attack could have the same economic effects, except that the buildings themselves would not be destroyed. In this case, the cyber attack may be less likely to be regarded as a use of force than a kinetic attack with the same (temporary) economic effect, simply because the lack of physical destruction would reduce the scale of the damage caused. However, a cyber attack against the stock exchanges that occurs repeatedly and continually, so that trading is disrupted for an extended period of time, for days or weeks, would surely constitute a use of force or even an armed attack, even if no buildings were destroyed.

- **A cyber attack is launched against the ground station of a Zendian military photo reconnaissance satellite**. Neither the

[17] *See* Barkham, *supra* note 15 (offering a related perspective).

satellite nor the ground station is physically damaged, but Zendia is temporarily unable to download imagery. The open question is whether such an act might plausibly be interpreted as a use of force, based on the argument that the inability to download imagery might be a prelude to an attack on Zendia, even if no permanent damage has been done to Zendia.

- **A cyber attack has effects that build slowly and gradually**. For example, a cyber attack against a stock exchange might corrupt the data used to make trades. Again, no physical damage occurs to buildings, and trading continues, albeit in a misinformed manner. Over time, the effects of such an attack could wreak havoc with the market.[18] If and when the effects were discovered, public confidence in the market could well plummet, and economic chaos could result. An open question is the degree of economic loss, chaos, and reduction in public confidence that would make such an attack a use of force.

- **A cyber attack is aimed at corrupting a manufacturing process**. In this scenario, the manufacturing process is altered in such a way that certain flaws are introduced into a product that do not show up on initial acceptance testing but manifest themselves many months later in the form of reduced reliability, occasional catastrophic failure, significant insurance losses, and a few deaths. Here, one open question relates to the significance of the effects of the attack. This is recognized as the "boiling the frog" phenomenon – a sudden change may be recognized as significant, but a gradual change of the same magnitude may not be.

2. *Definition of "Threat" of Force*

Article 2(4) of the U.N. Charter prohibits nations from threatening the use of force. When an actor wields traditional weapons as its coercive instruments, a threat generally takes the form of "We will do destructive act *X* if you do not take action *Y* or if you do take action *Z.*" That is, one actor is trying to compel its adversary to take action *Y* or deter its adversary from taking action *Z.* Does a threat to use existing vulnerabilities in an adversary

[18] As a demonstration that slowly accumulating error can have large consequences, consider that the Vancouver stock exchange index was introduced in 1982, and twenty-two months later, was undervalued by forty-eight percent compared to its "true" value. The reported value of the index was 524.881, whereas the correctly calculated value was 1009.811. This discrepancy was the result of round-off error accumulated over time. *See* B.D. McCullough & H.D. Vinod, *The Numerical Reliability of Econometric Software*, 37 J. ECON. LIT. 633 (1999).

computer system or network constitute a threat of the use of force under the U.N. Charter? Because an existing vulnerability can be used for cyber attack (which can be a use of force) or cyberexploitation, the answer is not clear. Does it matter how those vulnerabilities got there? Does introducing vulnerabilities into an adversary's system or network constitute a threat of force, especially if the vulnerabilities remain unused for the time being?

The following examples illustrate possible scenarios that raise questions about the definition of the threat of force:

- **Zendia introduces cyber vulnerabilities into the critical infrastructure of its adversary Ruritania but does not take advantage of them**. Since Ruritania suffers no ill effects from the fact that its infrastructure now has a number of vulnerabilities, no armed attack or even use of force has occurred. Ruritania learns of the Zendian penetration because its cybersecurity experts have detected it technically. Does the Zendian action of introducing cyber vulnerabilities constitute a threat of force against Ruritania? Does it make a difference if these vulnerabilities could be used equally well for cyberexploitation as for cyber attack? Does the possibility that Zendia could take advantage of those agents on a moment's notice make a cyber attack on Ruritania imminent, and if so, does it justify a Ruritanian strike on Zendia (cyber or otherwise) as an act of anticipatory self-defense?

  A helpful analogy is the idea of digging a tunnel underneath a border that terminates beneath a military facility. If Zendia digs such a tunnel under the Zendia-Ruritania border and Ruritania discovers it, Ruritania may well regard it as a hostile act. However, whether the tunnel amounts to an indication of imminent hostilities that would justify a Ruritarian strike on Zendia depends on many other factors.

- **Zendia discovers cyber vulnerabilities in the critical infrastructure of its adversary, Ruritania, but does not take advantage of them**. These vulnerabilities are found in software used by both Zendia and Ruritania and are supplied by a third-nation vendor. If Zendia notifies Ruritania of these vulnerabilities during a time of tension between the two nations, has Zendia threatened to use force against Ruritania?

3. Uncertainties in Identification and Attribution[19]

Application of *jus ad bellum* to cyber attacks requires identification of the party responsible for an act of cyber aggression because force must be directed against specific targets. No one has come close to solving the problem of technical attribution – the ability to identify the party responsible for an offensive cyber operation based only on technical indicators and information associated with that operation. This is not to say that attribution of an offensive cyber operation is impossible – for example, it may be possible to use information acquired from non-technical sources such as human intelligence to help ascertain the party responsible for the operation. But in the worst case, it may be difficult or impossible even to know when an offensive cyber operation has begun, who the attacker is, and what the operation's purpose and effects are or were. It may be very difficult to identify even the nature of the involved party, let alone the name of the country, terrorist group, or individual responsible.

The following examples illustrate possible scenarios in which uncertainties in identification and attribution arise:

- **During conflict between the United States and Zendia, the United States contemplates launching a cyber attack on a computer controlling a Zendian air defense network**. A normally reliable human informant passes on a message to the United States, but the message is unfortunately incomplete, and the only information passed along is the computer's electronic identifier, such as an IP address or a MAC address; its physical location is unknown. The open question is whether this computer is a valid military target for a U.S. cyber attack and the extent to which the United States has an obligation to ascertain its physical location prior to such an attack.

- **During a time of international tension (for example, U.S. forces are on an elevated alert status), the United States experiences a cyber attack on its military communications that is seriously disruptive**. The United States must restore its communications quickly but lacks the intelligence information to make a definitive assessment of the ultimate source of the attack. The open question is whether it can lawfully act against the proximate sources of the attack in order to terminate the threat and restore its communications capability, even though it is by no means certain that the "proximate source" is actually

[19] For an extensive discussion regarding the technical difficulties of attribution, see NRC Report, *supra* note 1, at §2.4.2, and NATIONAL RESEARCH COUNCIL, TOWARD A SAFER AND MORE SECURE CYBERSPACE (Seymour E. Goodman & Herbert S. Lin eds., 2007).

the ultimate source and may simply have been exploited by the ultimate source. (A proximate source might be a neutral nation or a nation whose relations with the United States are not particularly good. If the latter, a U.S. attempt to neutralize the attack might exacerbate tensions with that nation.)

4. The De Facto Exception for Espionage

If the traditional international legal regime regarding espionage is accepted, espionage conducted by or through the use of a computer – that is, cyberexploitation – is permissible, even if techniques are used that could also be used for destructive cyber attack. However, the distinction between a cyber attack and a cyberexploitation may be very hard to draw from a technical standpoint, since both start with taking advantage of a vulnerability.

Even in traditional espionage, espionage may raise LOAC issues if a clear distinction cannot be drawn between a given act of espionage and the use of force. For example, Roger Scott notes that certain forms of espionage involving ships, submarines, or aircraft as the collection platforms have been seen as military threats, and treated as matters of armed aggression permitting a military response, rather than domestic crimes demanding a law enforcement response.[20] One common thread in these cases appears to be that the collection platform is or could be a military asset such as a plane, a ship, or a submarine that could conduct kinetic actions against the targeted nation. The question of intent was central to the targeted nation at the time the potentially hostile platform was detected.

These issues are even more complicated in cyberspace. Recall that a payload of a software agent may have capabilities for both exploitation and destructive action, and that its capabilities may be upgradeable in real time. In this light, consider the following examples:

- **An offensive cyber operation introduces a two-part software agent into an adversary system**. The software agent is designed with two parts. One part is used for cyberexploitation, monitoring traffic through the system and passing the traffic along to a collection point. A second part is for cyber attack, awaiting an instruction to "detonate," at which point it will destroy the read-only memory controlling the boot sequence (process that starts the computer) of the machine where it resides. Until the agent detonates, no damage has been caused to the system, and no use of force has occurred. On the other hand, the potential to do damage has been planted and the

[20] *See* Roger D. Scott, *Territorially Intrusive Intelligence Collection and International Law*, 46 AIR FORCE L. REV. 217, 217 (1999).

act of planting the agent with a destructive component can be regarded as a threat of force. Under what circumstances, if any, does this offensive action constitute a use of force or the threat of force? The clandestine nature of the agent complicates matters further. An essential dimension of "threat" is that it must be known to the party being threatened, and there is a strong likelihood that the system owner does not know of the agent's existence. Still, the owner could discover the agent on its own and might well feel threatened after that point.

- **An offensive cyber operation introduces an upgradeable software agent into an adversary system**. As introduced, this agent conducts cyberexploitation, since it monitors traffic through the system and passes it along to a collection point. But through a software upgrade transmitted to the agent by clandestine means, the agent can then take destructive action, such as destroying the read-only memory controlling the boot sequence of the machine where it resides. A similar analysis applies in this instance. The agent as introduced does not constitute a use of force, as it has no destructive potential. But it can easily be turned into a destructive agent, and perhaps the act of upgrading the agent with a destructive component can be regarded as a threat of force or an imminent attack. Under what circumstances, if any, does this offensive action constitute a use of force or the threat of force?

- **A nondestructive probe is launched to map an adversary's computer network**. As such, this operation is a cyberexploitation. It is gathering intelligence on the network. Such an attack causes no damage to the network but provides the attacker with valuable information that can be used to support a subsequent cyber attack. An analogy might be drawn to the act of flying near an adversary's borders without violating its airspace in order to trigger radar coverage and then to gather intelligence on the technical operating characteristics of the adversary's air defense radars. Though such an act might well be regarded as unfriendly, it almost certainly does not count as a use of force.

5. Distinctions Between Economic Sanctions and Blockades[21]

Under international law, economic sanctions appear not to constitute a use of force, even if they result in death and destruction on a scale that would have constituted a use of force if they were caused by traditional military forces, although this interpretation is often questioned by the nation targeted by the sanctions. Article 41 of the U.N. Charter gives the Security Council authority to adopt measures "not involving the use of armed force," and explicitly recognizes that such measures include the "complete or partial interruption of economic relations."[22]

In this instance, international law does appear to differentiate between different means used to accomplish the same end. Economic sanctions and blockades could easily result in similar outcomes but have two key differences. First, sanctions are, by definition, a refusal of participating nations to trade with the targeted party, either unilaterally by virtue of a national choice or collectively by virtue of agreement to adhere to U.N. mandates regarding sanctions. Sanctions involve refraining from engaging in a nonobligatory trading relationship. By contrast, blockades interfere with trade involving any and all parties, willing and unwilling. Effective economic sanctions generally require coordinated multilateral actions, whereas blockades can be conducted unilaterally, though the coordination mechanism may or may not be tied to U.N. actions.[23]

Viewed from an effects-based analytical perspective, traditional LOAC thus has some inconsistencies as to its treatment of the means used for economic coercion – whether or not cyber attack is involved. At the very least, the LOAC does not draw entirely clear-cut distinctions. Accordingly, it is not surprising that inconsistencies might emerge if cyber attack is the means used for economic coercion, without immediate loss of life or property. Legal analysts must thus determine the appropriate analogy that should guide national thinking about cyber attacks that result in severe economic dislocation. In particular, are such cyber attacks like economic sanctions, or like a blockade, or even like some form of kinetic attack, such as the mining of a harbor?

[21] *See* Barkham, *supra* note 15 (providing an analysis that roughly parallels the argument of this subsection).

[22] U.N. Charter art. 41.

[23] Some economic sanctions can be imposed unilaterally and still be effective. For example, if the Zendian armed forces use a sophisticated weapons system that was originally produced in the United States, spare parts for that system may only be available from the United States. The United States could unilaterally choose to refrain from selling spare parts for that system to Zendia without violating the LOAC, and such an action could have significant effects on the Zendian armed forces as the weapons system deteriorated due to a lack of spare parts. In addition, multilateral sanctions need not necessarily involve the United Nations, as demonstrated by the Arab boycott of Israel, the Arab oil embargo of 1973, and the 2008 financial sanctions against Iran.

This question is particularly salient in the context of Internet-enabled commerce. The U.N. Security Council could decide to impose economic sanctions on a nation in order to compel that nation to follow some directive. Those sanctions could be quite broad. If a large part of the target nation's commerce was enabled through international Internet connections, the omission of such commerce from the sanctions regime might be a serious loophole.[24] Cyber attacks against the target nation might be required to prevent such commerce from taking place in a manner analogous to the U.N.'s use of naval and air forces to enforce certain past economic sanctions.

Two examples provided below illustrate possible scenarios that raise questions about whether to treat a cyber attack as a blockade or an economic sanction:

- **A continuing cyber attack effectively disconnects Zendia's access to the Internet, when Zendia is the target of U.N. economic sanctions**. In the modern era, the dependence of a nation's economic relations with the outside world on the Internet may be greater than the dependence of national economies on maritime shipping in the mid-twentieth century. Should this type of cyber attack – perhaps performed openly by a permanent member of the U.N. Security Council – be regarded as a blockade imposed through electronic means or as the enforcement of economic sanctions? Does it matter if the cyber attack targets only the Zendian connections to the outside world as opposed to targeting internal communications nodes and routers?

- **A cyber attack shuts down a key industry or segment of the armed forces of the targeted nation**. Economic sanctions and blockades can be narrowly tailored to affect only certain industries. For example, sanctions and blockades could prevent the sale or distribution of spare parts necessary for the continuing operation of a certain industry. The same is true for spare parts needed to maintain and operate certain weapons systems. A cyber attack could have similar effects and in particular could be carried out in such a way that the industry or military segment targeted was degraded slowly over time in a manner similar to its degradation due to the lack of spare parts. Thus, this kind of cyber attack could have effects

[24] As a practical matter, many of the nations that are subject to sanctions are often not heavily dependent on Internet commerce, or at least they are not today. In addition, sanctions are often not generalized but rather are targeted at specific goods such as arms.

> identical to that of either blockades or economic sanctions, though one is regarded as a use of force and the other is not.

IV. DISCUSSION

Offensive cyber operations pose a number of challenges to the U.S. commitment to abide by the LOAC and the U.N. Charter. These challenges are not irremediable, but they will require consideration of a number of factors that do not arise (or do not arise as strongly) when kinetic operations are concerned.

The first challenge is the technical similarity between cyber attack and cyberexploitation. Although cyber attack and cyberexploitation are conducted with very different intents, the latter can easily be mistaken for the former. This potential ambiguity has consequences both from the perspective of the targeted party (the adversary) and from that of the U.S. policy maker.

The party targeted in an offensive cyber operation is unlikely to know the intent behind the operation on the basis of the information that is available when the operation is discovered. Gaining access to take advantage of a vulnerability is how a cyber attack necessarily starts, and it is also how a cyberexploitation starts. Technical analysis of the payload may reveal its capability, but it cannot account fully for what the payload will actually do if it must wait for instructions, and it cannot account at all for what the payload will actually do if the payload can be remotely upgraded.

If the United States launches an offensive cyber operation, it cannot assume that the adversary will automatically understand the intent underlying it. We may believe that a cyber operation that penetrates an adversary's command and control system is conducted to gain intelligence information about the adversary's intent and operating procedures, but the adversary may well believe that such a penetration demonstrates hostile intent of the United States because it cannot be sure that we have not also compromised the operational effectiveness of their command and control system.

Thus, it may make sense for the United States to take steps to increase the likelihood that the adversary will not misinterpret U.S. actions. For example, the United States might choose to conduct its offensive cyber operations in such a way that cyberexploitations are clearly distinguishable in a technical sense from cyber attack. Such a choice would reduce the operational flexibility of its instruments for offensive cyber operations by restricting the capabilities of an already inserted agent, but would minimize the likelihood that an adversary would mistake a cyberexploitation for a cyber attack.

Making such a choice is "above the pay grade" of field operators. Such a choice should be the responsibility of senior policy makers. On the other hand, senior policy makers generally do not concern themselves with operational details of a mission, and because cyberexploitation and cyber attack share so many technical and operational similarities, it is easy to imagine that dedicated and professional field operators would as a matter of course equip a cyber instrument with both exploitation and attack capabilities (or with the capability for being upgraded) unless they were specifically instructed not to do so. Similarly, it is easy to imagine that because of the relatively small operational footprint of an offensive cyber operation, senior policy makers will need special mechanisms put into place to notify them of any such operations that might be planned, especially during a crisis, when their initiation might be unduly provocative or otherwise inappropriate.

A second challenge involves the inevitable uncertainties associated with understanding the nature and scope of an offensive cyber operation. Even if it is known that a given operation is an attack or an exploitation, and leaving aside the question of attribution, determining the overall scale of the effects may well require an analysis and correlation of events at multiple sites. When it is discovered that something is happening, the target of an offensive cyber operation is not likely to be able to distinguish between an offensive cyber operation that seeks to cause large-scale damage (a cyber attack that would almost certainly constitute an armed attack) and one that seeks to cause only very limited damage (a cyber attack that might constitute a use of force but not an armed attack).

The challenge faced by a nation trying to figure out whether a given act that may appear to be hostile is a precursor to more serious hostile actions that will create additional damage is not unique to cyber attacks, as illustrated by the Tonkin Gulf incident (in which the United States was arguably too quick to see a grave provocation) and Stalin's refusal to believe reports of Nazi preparations and initial incursions in June 1941. An aircraft penetrating a nation's airspace without authorization may simply be off course, or it may be carrying nuclear weapons with hostile intent. The nation in question has an obligation to try to determine if the aircraft represents a true threat, but the nation surely has a right to shoot down the craft if it reasonably concludes that it is at risk. The open question is what the nation can do if it is uncertain about whether the aircraft poses a threat.

Definitive answers to questions such as attribution and scope are likely to be unavailable immediately after an offensive cyber operation is detected. Although waiting to see what course the operation takes is the only certain way to determine the scale and extent of its effects, waiting may not be a viable option for decisionmakers who believe that a cyber attack on their nation is underway. In addition, leaders of a state often wish to calibrate a response to an attack. If decisionmakers do not know the scale of the attack, how are they to calibrate a response?[25] Thus, it is likely

[25] *See* Barkham, *supra* note 15, at 80-83.

that decisionmaking under these circumstances will have to take place under conditions of great uncertainty and intense time pressure.

A third challenge arises from the traditional relationship between espionage and international law. Although espionage has not traditionally been regarded as a use of force, some analysts argue that cyberexploitations against sensitive military or intelligence sites conducted over an extended period and in large volume constitute a demonstration of hostile intent that may indeed violate U.N. Charter provisions prohibiting the use of force.

Applying this rule to the cyber domain raises the question of what actions constitute a demonstration of hostile intent. For example, do nondestructive probes of important military U.S. computer systems and networks or even systems and networks associated with U.S. critical infrastructure constitute demonstrations of hostile intent? If so, do such actions justify responses beyond the taking of additional passive defense measures? Would a commander be permitted to conduct probes of adversary networks from which these probes were emanating? To conduct a responsive cyber attack to neutralize the probes?

The presence of uncertainties in understanding the nature and scope of an offensive cyber operation means that a nation seeking U.N. action in response to a cyber attack would be unlikely to see rapid action, because much of the necessary information might not be immediately available. Indeed, one might consider as a benchmark the history of extended Security Council debate on authorizations for armed conflict involving kinetic force.

The fourth and last challenge raised in this article is the inconsistency between economic sanctions, avowedly not a use of force and thus an entirely permissible unilateral action under the U.N. Charter, and blockades, avowedly a use of force and thus a violation of Article 2(4) unless authorized by the U.N. Security Council or undertaken as a national exercise of its inherent right to self-defense. This inconsistency – present even if cyber operations are not at issue – is even more problematic when cyber-assisted sanctions and blockades become possible courses of action.

CONCLUSION

Offensive cyber operations (cyber attacks and cyberexploitations) pose many challenges for the interpretation of *jus ad bellum* and the U.N. Charter, some of which have been addressed in this article. In the long run, it is inevitable that some future conflict will have a cyber component to it, and it behooves policy makers to understand the legal landscape before such a conflict occurs. A central recommendation of the NRC Report is that analysts develop the requisite knowledge and expertise now so that they are prepared to help policy makers if and when such conflict occurs. By exploring some of the relevant questions, this article takes one modest step on that path of exploration.

APPENDIX

Cyber Attack versus Cyberexploitation

Terms	**Cyber attack, computer network attack**	**Cyberexploitation, intelligence exploitation, computer network exploitation**
Approach and intent	Degrade, disrupt, deny, destroy attacked infrastructure and systems/networks	Conduct smallest intervention consistent with desired operations
Primary relevant domestic law	U.S. Code Title 10 authorities and restrictions*	U.S. Code Title 50 authorities and restrictions
Operational agency	U.S. Strategic Command, Joint Functional Combatant Command for Network Warfare (will become U.S. Cyber Command)	National Security Agency
Main advocate in the U.S. government to date	U.S. Air Force	Director of National Intelligence
Interactions with tactical military operations	Based on explicit inclusion in battle plans	Based on intelligence reporting
Characterization of personnel	War fighters	Intelligence community

* Covert action involving cyber attack would fall under Title 50 authorities.

Illustrative Vulnerabilities of Systems and Networks

Vulnerabilities	Description
Software	Application or system software may have accidentally or deliberately introduced flaws the use of which can subvert the intended purpose for which the software is designed.
Hardware	Vulnerabilities can be found in hardware, including microprocessors, microcontrollers, circuit boards, power supplies, peripherals such as printers or scanners, storage devices, and communications equipment such as network cards. Tampering with such components may secretly alter the intended functionality of the component or provide opportunities to introduce hostile software.
Seams between hardware and software	An example of such a seam might be the reprogrammable read-only memory of a computer (firmware) that can be improperly and clandestinely reprogrammed.
Communications channels	The communications channels between a system or network and the "outside" world can be used by an adversary in many ways. An adversary can pretend to be an authorized user of the channel, jam it, and thus deny its use to the adversary, or eavesdrop on the channel to obtain information intended by the adversary to be confidential.
Configuration	Most systems provide a variety of configuration options that users can set based on their own tradeoffs between security and convenience. Because convenience is often valued more than security, many systems are – in practice – configured insecurely.
Users and operators	Authorized users and operators of a system or network can be tricked or blackmailed into doing the bidding of an adversary.
Service providers	Many computer installations rely on outside parties to provide computer-related services, such as maintenance or Internet service. An adversary may be able to persuade a service provider to take some special action on its behalf, such as installing attack software on a target computer.