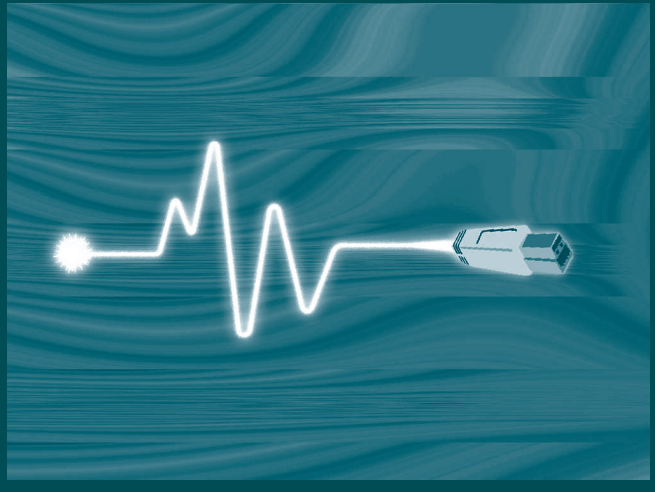
United Nations
Institute for
Disarmament Research
UNIDIR

disarmament forum

four • 2011



Confronting cyberconflict

The United Nations Institute for Disarmament Research is a voluntarily funded, autonomous institution within the framework of the United Nations.

Through its research projects, publications, conferences and expert networks, UNIDIR promotes creative thinking and dialogue on both current and emerging security challenges.

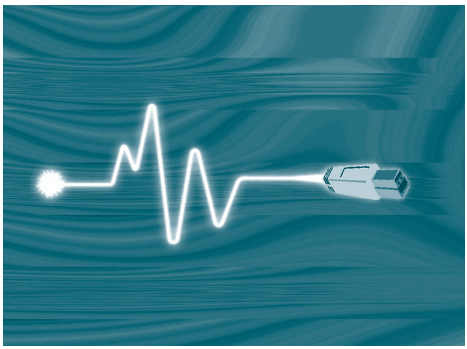
Disarmament Forum is supported by the contributions of the Governments of Finland, France, Hungary, Indonesia, Ireland, Israel, Luxembourg, Malaysia, Mexico, Norway, Pakistan, the Russian Federation, Switzerland and Turkey..

UNIDIR—ideas for peace and security www.unidir.org

Cover design by Diego Oyarzún Reyes

disarmament forum

four • 2011



Confronting cyberconflict

Palais des Nations
CH-1211, Geneva 10, Switzerland
Tel.: +41 (0)22 917 31 86
Fax: +41 (0)22 917 01 76
disarmamentforum@unog.ch
www.unidir.org

United Nations
Institute for
Disarmament Research
UNIDIR

Printed at United Nations, Geneva
GE.12-00703—May 2012—4,270
UNIDIR/2012/1
ISSN 1020-7287

Printed on recycled paper

Table of contents

Editor's note

Kerstin Vignard

The spectre of cyberconflict has finally captured the world's attention and imagination—from the highest levels of governments, to the covers of magazines, to the scripts of both Hollywood and Bollywood. Putting the doomsday hype aside, the fact that new technologies are exploited for offensive and defensive purposes is nothing new. Cyberconflict is simply conflict carried out with the latest "weapons" humanity has at hand. What is challenging about the issue of cyberconflict is that it exploits, in many cases, widely used dual-use technologies such as computer networks and the Internet itself, and that the number of potential actors—governments, hackers, terrorists, the private sector, criminals, and even unwitting computer users—has exponentially grown.

Technological development often rushes ahead of legal, definitional and ethical debates—and cyber development is no different. The international community is now starting the process of discussion with the goal of reaching common understandings. This issue of *Disarmament Forum* is a contribution to that critical discussion.

The next issue of *Disarmament Forum* will look ahead to the 2013 Chemical Weapons Convention Review Conference, and considers some of the remaining and emerging challenges to the CW regime. The rapid pace of scientific and technological developments means that the chemical weapons regime must be agile, forward looking and practical in nature. Are states parties to the CW regime prepared to address remaining treaty ambiguities, such as those relating to incapacitating chemical agents? What are the consequences of the fact that Libya, the Russian Federation and the United States have all indicated that they will not be able to meet the April 2012 deadline for the destruction of declared chemical weapons? As the verification of destruction activities will be significantly reduced in the coming years, what will be the key roles and functions of the Organization for the Prohibition of Chemical Weapons? What impact will these shifting priorities have on its organizational structure? Last but not least, how will a new focus on preventing the re-emergence of chemical weapons relate to the goals of international cooperation and assistance? How will this be embedded in strengthened national implementation measures more broadly?

UNIDIR's annual space security conference took place from 29–30 March. Entitled "Laying the Groundwork for Progress", the 2012 conference provided a platform for building understanding and facilitating discussion on pressing issues affecting stability in outer space. More information on this and previous conferences, including reports and audio files, can be found on our website.

Over 50 heads of state and international organizations recently met in Seoul, Republic of Korea, for the 2012 Nuclear Security Summit. As a contribution to the summit preparations, UNIDIR produced *Global Nuclear Security: Building Greater Accountability and Cooperation*. The book provides an overview of the international agreements, programmes and institutional

arrangements that form the core of the international nuclear security regime. Full details of the publication can be found in the *UNIDIR focus* of this issue and on our website.

2012 has begun as a year of transitions. After four years of service as UNIDIR Deputy Director, Christiane Agboton-Johnson left the Institute at the end of February. Christiane brought her experience and passion to bear on both UNIDIR's programme of work and its internal processes. And with this issue, *Disarmament Forum* says farewell to our English editor, Ross McRae. Ross's enthusiasm and initiative added a new dimension to the team. Valérie and I, on behalf of all of our UNIDIR colleagues, wish both of them the very best in their next endeavours.

Cyber operations and *jus in bello*

Nils Melzer

International humanitarian law (IHL), sometimes also described as the "law of armed conflict" or *jus in bello*, applies exclusively in situations of armed conflict and regulates the conduct of hostilities between the belligerent parties, as well as the protection and treatment of those having fallen into the power of the enemy.[1] Today, the most important sources of IHL are the four Geneva Conventions of 1949 and the first two Additional Protocols of 1977 (Protocols I and II), as well as the Regulations concerning the Laws and Customs of War on Land revised at the Hague Conference of 1907, and a series of treaties prohibiting or restricting the use of certain weapons. Additionally, in the course of decades and centuries of warfare, a rich body of customary IHL has developed, which proves helpful in cases not regulated by applicable treaty law.[2]

Cyber operations as warfare

The notions of "cyberwar", "cyberwarfare", "cyberhostilities" and "cyberconflict" have not been authoritatively defined for the purposes of international law. The only treaty definition that exists is from the Shanghai Cooperation Organization and concerns the wider concept of "information war", which is defined as confrontation between two or more states in the information space aimed at damaging information systems, processes and resources, and undermining political, economic and social systems, mass brainwashing to destablize society and state, as well as forcing the state to take decisions in the interest of an opposing party.[3]

As pointed out by Michael Schmitt, a leading commentator, the term "information warfare" is often inaccurately used as a synonym for "information operations": while the latter can occur both in times of peace and of war, the former refers exclusively to information operations conducted in situations of armed conflict and excludes information operations occurring during peacetime.[4]

Applied to the more specific context of cyber operations, this means that the use of the terms "cyberwar", "cyberwarfare", "cyberhostilities" and "cyberconflict" should be restricted to armed conflicts within the meaning of IHL. Indeed, security threats emanating from cyberspace which do not reach the threshold of armed conflict can be described as "cybercrime", "cyber operations", "cyberpolicing" or, where appropriate, as "cyberterrorism" or "cyberpiracy",

---

Nils Melzer is Research Director of the Competence Centre for Human Rights at the University of Zurich, and former Legal Adviser to the International Committee of the Red Cross (ICRC). He is currently a participating expert in a process sponsored by the North Atlantic Treaty Organization (NATO) Cooperative Cyber Defence Centre of Excellence to draft a manual on international law applicable to cyberconflict. The opinions expressed in this article are the author's own and do not necessarily reflect the views of the University of Zurich, the ICRC, NATO or the United Nations. The article was originally written in English.

but should not be referred to with terminology inviting doubt and uncertainty as to the applicability of the law of armed conflict.

Cyber operations in current conflicts

Today it appears to be uncontested that IHL applies to cyber operations which are carried out in the context of an ongoing international or non-international armed conflict.[5] It seems to be generally recognized that the non-existence of cyber operations at the time when most contemporary instruments of IHL were drafted and adopted does not today preclude their applicability to such operations. One of the most fundamental rules of IHL is the "right of belligerents to adopt means of injuring the enemy is not unlimited",[6] and Article 36 of the Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I) expressly requires that:

> In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.

Existing IHL clearly anticipates the application of its rules and principles to newly developed methods and means of warfare. It is not the precise nature of a means or method, but the context in which it is used which subjects it to the rules and principles of IHL. Whether a cyber operation must be regarded as carried out in the context of an armed conflict does not necessarily depend on the territorial connection of the operation but rather on whether it is carried out for reasons related to an armed conflict or, in the words of the International Criminal Tribunal for the Former Yugoslavia, whether it has a nexus with an ongoing armed conflict.[7] This also means that cyber operations conducted for reasons unrelated to an armed conflict (lack of nexus) may qualify, for example, as cybercriminality or cyberpolicing, but are not governed by IHL—even if carried out by a belligerent party or within a territory affected by an armed conflict.

Can cyber operations trigger an armed conflict?

One of the most difficult questions is whether and under which circumstances cyber operations can give rise to an armed conflict and thus trigger the applicability of IHL without the parallel occurrence of conventional hostilities. This question must not be confused with the distinct questions of whether cyber operations can qualify as a threat or use of force or an armed attack within the meaning of the Charter of the United Nations. According to the International Committee of the Red Cross (ICRC), currently prevailing legal opinion on the definition of armed conflict under IHL can be summarized as follows:

1. **International armed conflicts** exist whenever there is *resort to armed force between two or more States*.
2. **Non-international armed conflicts** are *protracted armed confrontations* occurring between governmental armed forces and the forces of one or more armed groups, or between such groups arising on the territory of a State [party to the Geneva Conventions]. The armed confrontation must reach a *minimum level of intensity* and the parties involved in the conflict must show a *minimum of organisation*.[8]

While a situation can evolve from a non-international to an international armed conflict and vice versa, the ICRC states: "Legally speaking, no other type of armed conflict exists".[9] Thus, cyber operations can trigger the applicability of IHL if they give rise to all required constitutive elements of an international or non-international armed conflict.

As far as international armed conflicts are concerned, cyber operations must amount to the "resort to armed force between two or more States". The question of whether armed force occurs "between" states essentially turns on legal attributability as governed by the general international law of state responsibility. Accordingly, the applicability of IHL cannot be limited to acts committed by members of the state armed forces, but must be extended to the conduct of any other person acting as a state agent, whether *de jure* or *de facto*, on behalf of a belligerent. While there is no reason to alter the application of the law of state responsibility in cyberspace, the identification of the source or author of a cyber operation can be particularly difficult.

The second question is whether cyber operations can be regarded as armed force (or, in non-international armed conflict, as armed confrontation) triggering the applicability of IHL even in the absence of kinetic force. So far, there seems to be consensus that this is the case—at least wherever cyber operations cause the same effects as kinetic force—namely death, injury or destruction.[10] However, not every use of force indicates the existence of an armed conflict and not all acts of war necessarily involve a use of force. Indeed, armed conflicts can even be triggered by formal declarations of war. Strictly speaking, therefore, the existence of an international armed conflict does not necessarily depend on the use of force between states but—at least in the absence of a formal declaration of war—on the occurrence of belligerent hostilities within the meaning of IHL. Accordingly, state-sponsored cyber operations would give rise to an international armed conflict if they are designed to harm another state, not only by directly causing death, injury or destruction, but also by directly adversely affecting its military operations or military capacity.

Non-international armed conflict differs in that it involves at least one non-state belligerent showing a minimum degree of organization and armed confrontations or hostilities must show a minimum level of intensity. The first criterion requires organized collective action, which would certainly exclude cyber operations conducted by individual hackers from the notion

of armed conflict. From a strictly theoretical perspective it cannot be excluded that even a small but organized group of hackers launching highly destructive cyber operations against a state's military network could trigger a non-international armed conflict. As long as such cyber operations emanate from within territory controlled by the state under attack, however, and as long as they are not accompanied by a threat or use of conventional military force, which could prevent the state from exercising its territorial authority over the attackers, such operations would most likely be regarded as a criminal threat to be addressed through law enforcement measures. A qualification of such operations as "hostilities" capable of triggering a non-international armed conflict becomes more likely when they occur repeatedly and emanate from territory where the attacked state cannot exercise its law enforcement authority, and where the local authority is unwilling or unable to intervene.

It is probably still too early to make definite statements on the precise threshold at which cyber operations trigger a non-international armed conflict (a question unresolved even for non-international conflicts fought through traditional means and methods). As has rightly been stated in the ICRC contribution to the 2004 Stockholm Conference, "[w]hether CNA [computer network attacks] alone will ever be seen as amounting to an armed conflict will probably be determined in a definite manner only through future state practice".[11]

In any case, once the existence of an armed conflict has been established, it will have to be determined to what extent traditional concepts and rules of IHL can be transposed to cyber operations conducted in the context of that conflict. This article will focus on examining those concepts and principles which are likely to be most relevant in practice, namely the concepts of "attack", "hostilities" and "direct participation", as well as the rules and principles governing targeting and good faith in the conduct of hostilities.

Cyber operations as attacks

The term "attack" is an important technical term of IHL in that many of its fundamental rules on the conduct of hostilities are expressed in terms of attacks. For example: "the civilian population as such, as well as individual civilians, shall not be the object of attack";[12] "civilian objects shall not be the object of attack";[13] "indiscriminate attacks are prohibited";[14] and "attacks shall be limited strictly to military objectives".[15] The same applies, inter alia, to the rules regulating "precautions in attack" and "precautions against the effects of attack",[16] those protecting medical units,[17] persons *hors de combat*,[18] works and installations containing dangerous forces[19] against attack, as well as those obliging combatants to distinguish themselves from the civilian population during attack or military operations preparatory to an attack,[20] and those prohibiting the use of the flags or military emblems, insignia or uniforms of adverse parties during attack.[21]

According to Article 49(1) of Protocol I: "'Attacks' means acts of violence against the adversary, whether in offence or in defence". This definition has triggered significant discussion as to

what extent cyber operations, in view of their non-kinetic nature, could be regarded as acts of violence and therefore as attacks within the meaning of IHL. It is generally recognized that acts of violence do not necessarily require the use of kinetic violence, but that it is sufficient if the resulting effects are equivalent to those normally associated with kinetic violence—namely the death or injury of persons or the physical destruction of objects (effects-based approach). Strictly speaking, this approach does not extend the notion of attack beyond acts of violence, but simply recognizes that cyber operations triggering processes likely to directly cause death, injury or destruction are not only equivalent to but constitute an integral part of an act of violence within the meaning of Article 49(1).[22]

There is disagreement, however, as to whether the notion of attack also includes cyber operations aiming to merely capture or neutralize (that is, inhibit, hinder or hamper the proper exercise of its function) rather than kill, injure or destroy the target. The leading argument in favour of extending the effects-based interpretation of attack to cyber operations aiming to neutralize is that the treaty definition of military objectives in Article 52(2) of Protocol I includes objects whose "capture or neutralization" would offer a definite military advantage and puts these two alternatives on the same level as "total or partial destruction".[23] Those opposing this extension base themselves on a more literal interpretation of attacks as "acts of violence" and require that, if not the act itself, at least its consequences must be violent in order for it to be considered an attack.[24] In support of this view they further point out that the principle of proportionality is formulated in terms of attacks "which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof" but does not include capture or neutralization.[25]

While both arguments have strong points, neither seems to provide an entirely satisfactory interpretation of the notion of attack in relation to cyber operations. On the one hand, it would hardly be convincing to exclude the non-destructive incapacitation of a state's air defence system or other critical military infrastructure from the notion of attack simply because it does not directly cause death, injury or destruction. On the other hand, it may well be exaggerated to extend the notion of attack to any denial of service attack against, for example, online shopping services, travel agents or telephone directories.

Although the term "attack" is a key notion of IHL, an analysis of the relevance of its rules on the conduct of hostilities for cyber operations cannot be limited to an examination of this notion. Recall, for example, that the basic treaty rule of distinction is not formulated in terms of "attacks" but in terms of "operations".[26] Similarly, treaty law protects the civilian population not only from direct attacks, but more generally from the "dangers arising from military operations"[27] and requires that "[i]n the conduct of military operations, constant care shall be taken to spare the civilian population, civilians and civilian objects".[28] Also, at least for states party to Protocol I, the prohibition of perfidy applies not only for operations aiming to injure or kill, but also to those aiming to capture an adversary.[29] Most persuasive, however, is the fact that civilians lose their protection "for such time as they take a direct part in hostilities",[30] a notion that is

generally considered to be wider than that of attack.[31] Therefore, although attacks certainly represent the predominant form of combat operation, it would be inaccurate to assume that cyber operations not amounting to an attack are not subject to IHL governing the conduct of hostilities. Accurately understood, the applicability of the restraints imposed by IHL on the conduct of hostilities to cyber operations does not depend on whether the operations in question qualify as "attacks" (that is, the predominant form of conducting hostilities), but on whether they constitute part of the "hostilities" within the meaning of IHL.

Cyber operations as hostilities and direct participation therein

According to the ICRC, hostilities refer to the (collective) resort by the parties to the conflict to means and methods of injuring the enemy and can be described as the sum of all hostile acts carried out by individuals directly participating in those acts.[32] In IHL the notion of "direct participation in hostilities" also describes the conduct which, if carried out by civilians, entails the suspension of their protection against direct attack.[33] Thus, for such time as civilian experts or individual hackers carry out cyber operations amounting to direct participation in hostilities, they are not only bound to comply with IHL governing the conduct of hostilities, but also become legitimate military targets—just as if they were combatants. Moreover, civilians directly participating in hostilities do not have to be taken into account when taking precautions in attack, most notably with a view to avoiding or minimizing incidental harm (so-called "collateral damage").

According to the ICRC's official position, the notion of direct participation in hostilities goes beyond the notion of attack and includes not only the infliction of death, injury or destruction, but essentially any act likely to adversely affect the military operations or military capacity of a belligerent party (*threshold of harm*).[34] Additionally, in order for a cyber operation to be considered part of the hostilities, it must cause the required threshold of harm directly (*direct causation*), and it must also be designed to do so in support of a belligerent and to the detriment of another (*belligerent nexus*). Whether the causal link between a specific operation and the resulting harm is "direct" or "indirect" depends, in essence, on whether it merely builds up the capacity of a belligerent party to harm the enemy (indirect) or whether it is an integral part of an operation using such capacity to actually inflict harm on the enemy (direct). Accordingly, where cyber operations attributable to a belligerent party are designed to harm the adversary, either by directly causing death, injury or destruction, or by directly adversely affecting military operations or military capacity, such operations must be regarded as "hostilities" and, therefore, are subject to all restrictions imposed by IHL on the choice and use of means and methods of warfare. If conducted by civilians, such operations also entail loss of protection against direct attacks.

Cyber operations aiming to disrupt or incapacitate an adversary's radar or weapons systems, logistic supply or communication networks may not directly cause any physical damage, but would certainly qualify as part of the hostilities and, therefore, would have to comply with the

rules and principles of IHL governing the conduct of hostilities.[35] The same would apply to cyber operations intruding into the adversary's computer network to delete targeting data, manipulate military orders, or change, encrypt, exploit or render useless any other sensitive data with a direct (adverse) impact on the belligerent party's capacity to conduct hostilities. However, cyber operations causing neither death, injury or destruction, nor military harm—such as those conducted for the purposes of general intelligence gathering, for purely criminal purposes or otherwise unrelated to the hostilities—would fall short of the concept of hostilities and, if conducted by civilians, would not entail loss of protection against direct attacks.

The most difficult question that remains unresolved in this respect is whether destruction necessarily presupposes physical damage, particularly in the absence of military harm. In other words, while the non-destructive incapacitation of a military computer network would clearly amount to military harm and thus automatically also to hostilities, the non-destructive incapacitation of a power station used exclusively for civilian purposes would cause neither military harm nor death, injury or destruction—unless destruction is interpreted as including harm other than physical damage. Again, this results in a dilemma between adopting either a too restrictive or a too permissive interpretation of the law. In the first case, even cyber operations causing the incapacitation of major civilian electric grids and communication networks could only qualify as part of the hostilities where they result in death, injury, physical destruction or military harm. In the second case, essentially any harm caused to the civilian population for reasons related to the conflict, including mere harassment or inconvenience, would have to be regarded as part of military hostilities. This would trigger not only the applicability of IHL on the conduct of hostilities, but also the loss of civilian protection for all those directly involved.

Targeting in cyberspace

At the heart of IHL lies the principle of distinction, which requires belligerent parties to always distinguish between legitimate military targets and persons and objects protected against attack, and to direct their operations only against the former.[36] Derived from the principle of distinction, and indispensable for its faithful implementation, are the prohibition of indiscriminate attack and the requirements of precaution and proportionality.

Persons

As far as persons are concerned, legitimate military targets include combatants, members of organized armed groups and civilians directly participating in hostilities. Civilians, medical and religious personnel, and combatants *hors de combat*—due to wounds, sickness, capture, surrender or any other reason—must be spared and protected. Although the identification of decisive factors such as direct participation in hostilities and membership in irregularly constituted armed forces or groups can pose significant practical difficulties, most of these

problems are not cyber-specific and have been discussed in more detail elsewhere.[37] Of particular relevance is, however, the question of how targeting-relevant factors, such as a group's organization or membership, should be interpreted in cyberspace, where persons may act collectively, without lasting affiliation or hierarchical command structure. In addition, how does the obligation of combatants "to distinguish themselves from the civilian population while they are engaged in an attack or in a military operation preparatory to an attack"[38] play out in cyberspace? Does it require hackers to wear uniforms even when far removed from the physical battlefield, or does it mean that their operations have to be recognizable as military operations to the adversary? How does this obligation relate to the distinction between (permitted) ruses of war and (prohibited) perfidy on the battlefield? It is clear that these and other questions need urgent clarification if civilians exposed to cyberwarfare are to receive the protection they are entitled to under treaty and customary law. In the meantime, it may have to suffice to recall that, in case of doubt, any person must be presumed to be a civilian and, as such, protected against direct attack.[39]

Objects

According to Article 52(2) of Protocol I:

> In so far as objects are concerned, military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offers a definite military advantage.

The challenge lies with the concrete implementation of this definition in cyberspace, which relies heavily on civilian infrastructure tightly interconnected with military cyber infrastructure. Even more than in traditional warfare, military objectives in cyberspace are likely to be dual use. While this does not represent an absolute obstacle against attacking such objects, it requires a high level of precaution in identifying legitimate targets, as well as a comparatively sophisticated capability of both the attacker and the attacked for assessing, avoiding and controlling incidental harm likely to be inflicted on the civilian infrastructure and population.

In view of the prohibition of indiscriminate attacks, the question arises to what extent malware intended to damage military systems can be prevented from spreading to civilian infrastructure and causing havoc among the civilian population.[40] Even if the collateral effects can be controlled it may be asked to what extent it would be justified, for example, to incapacitate a domain name server directing global internet traffic or destroy a major intercontinental submarine cable, in order to prevent their use for hostile cyber operations if more than 90% of the data transmitted are of civilian nature and the consequences for global trade, traffic and communication would be debilitating.[41]

Another key issue to be resolved is whether data constitutes an object within the meaning of IHL, and if so, what threshold of damage, modification, manipulation or interference would be required for the prohibition of attacks against civilian objects to be violated. Virtually no cyber operation—not even espionage through computer network exploitation or manipulations as simple as entering a password—can be carried out without at least temporarily deleting or changing data in the intruded systems. For the purposes of targeting, data should probably be regarded as an object which may not be directly targeted unless it fulfils all defining elements of a military objective.[42] The unavoidable (but incidental) deletion or modification of civilian data in the course of an operation pursuing a different aim, on the other hand, must be factored into the proportionality assessment, where the nature of the inflicted harm can duly be taken into account.

It is therefore important to distinguish the actual aim of the operation from its incidental side effects. For example, the deletion or modification of civilian data in the course of an attack against military cyber infrastructure would be equivalent to the kinetic causation of so-called "collateral damage". The manipulation or modification of access data to a civilian computer system in the course of an espionage or reconnaissance operation, on the other hand, could perhaps be compared to breaking the door or mailbox of a civilian house in the course of a search operation—but it would not constitute an "attack" within the meaning of Article 49 of Protocol I because neither the nature and effects nor the aim of the operation as such is equivalent to that of an "act of violence".[43] More difficult are examples such as the deletion or manipulation of data to disrupt civilian television broadcasts, which may be regarded as lawful by some,[44] whereas others would likely condemn it as a direct attack against a civilian object.[45]

Ruses and perfidy in cyberspace

The specific characteristics of cyberspace invite a plethora of opportunities and techniques to deceive the enemy with false information. Belligerents can disguise the origin of their operations through botnets and IP spoofing,[46] electronically camouflage combat troops or vehicles as medical transports, manipulate the enemy's reconnaissance data, and even send seemingly innocent emails infected with malware to military headquarters.

An important distinction must be made between (permitted) ruses of war and (prohibited) perfidy. Article 37 of Protocol I defines ruses of war as "acts which are intended to mislead an adversary or to induce him to act recklessly but which infringe no rule of international law applicable in armed conflict". Prohibited perfidy, on the other hand, refers to the killing, injuring or capturing of an adversary by leading him to believe that he is entitled to, or is obliged to accord, IHL protection, and subsequently betraying that confidence. Examples of perfidy include:

- the feigning of an intent to negotiate under a flag of truce or of a surrender;
- the feigning of an incapacitation by wounds or sickness;

- the feigning of civilian, non-combatant status; and
- the feigning of protected status by the use of signs, emblems or uniforms of the United Nations or of neutral or other states not parties to the conflict.[47]

However, IHL prohibits the resort to perfidy only in connection with the killing, injuring or capturing of an adversary. Cyber operations limited to causing physical or functional damage to infrastructure and other forms of disruption or incapacitation, even if conducted by resort to perfidious deception, would not come under this prohibition.

More relevant for cyber operations are the broader prohibitions on misusing internationally recognized protective emblems (for example, those used by the Red Cross and Red Crescent Movement, the flag of truce and the protective emblem of cultural property), the emblem of the United Nations, and the flags or military emblems, insignia or uniforms of neutral or states not parties to the conflict. It is also prohibited to use the flags or military emblems, insignia or uniforms of the adversary while engaging in attacks or in order to shield, favour, protect or impede military operations.[48] This would clearly outlaw any hostile cyber operation pretending to originate from a non-belligerent state, the ICRC or the United Nations, as well as attacks disguising themselves as operations conducted by friendly forces.

The status of cyberwarriors

Combatants

Cyber operations are generally carried out by highly specialized personnel. To the extent that they are members of the armed forces of a belligerent state, their status, rights and obligations are no different from those of traditional combatants. As laid out in Article 43 of Protocol I, the armed forces are defined as "all organized armed forces, groups and units which are under a command responsible to that Party for the conduct of its subordinates". This broad and functional concept of armed forces includes essentially all armed actors belonging to a belligerent state and showing a sufficient degree of military organization.

Contractors and civilian employees

Belligerent states have increasingly been employing private contractors and civilian employees in a variety of functions traditionally performed by military personnel—including the support, preparation and conduct of cyber operations. As long as such personnel assume functions not amounting to direct participation in hostilities, they remain civilians and, if formally authorized to accompany the armed forces in an international armed conflict, are even entitled to prisoner-of-war status in the case of capture according to Article 4 of the Geneva Convention relative to the Treatment of Prisoners of War. However, where private contractors or civilian employees are expressly authorized by a state to directly participate in hostilities on its behalf, they become organized armed actors and, *de facto*, irregular members of its armed forces.

As such, they lose civilian status and are entitled to combatant privilege and prisoner-of-war status as long as they fulfil the so-called "four requirements" restated in Article 4:

- being commanded by a person responsible for subordinates;
- having a fixed distinctive sign recognizable at a distance;
- carrying arms openly; and
- conducting their operations in accordance with the laws and customs of war.

Levée en masse

The term *levée en masse* refers to the inhabitants of a non-occupied territory who, on the approach of the enemy, spontaneously take up arms to resist the invading forces without having had time to form themselves into regular armed units, provided they carry arms openly and respect the laws and customs of war. Participants in a *levée en masse* are the only armed actors who are entitled not only to prisoner-of-war status, but also to the combatant privilege although, by definition, they operate spontaneously and lack sufficient organization and command to qualify as members of the armed forces. While this category of persons has become ever less relevant in traditional warfare, it may well come to be of practical importance in cyberwarfare, where territory is neither invaded nor occupied—which may significantly prolong the period during which a *levée en masse* can operate. Cyberspace also provides an ideal environment for the instigation and non-hierarchical coordination of spontaneous, collective and unorganized cyber defence action by great numbers of "hacktivists". The only question is, of course, how the requirement to "carry their arms openly" should be interpreted in cyberspace. From a teleological perspective, a possible solution would be to consider this requirement as fulfilled when cyber operations are not conducted by feigning protected, non-combatant status within the meaning of the prohibition of perfidy.[49]

Civilians

In IHL the concept of civilian encompasses all persons who are neither members of the armed forces of a state or non-state party to an armed conflict, nor participants in a *levée en masse*. As civilians, they are entitled to protection against the dangers arising from military operations and, most notably, against attack. In cyberwarfare this category is likely to include most non-state hackers not belonging to the military wing of an organized armed group. If and for such time as their operations amount to direct participation in hostilities, civilians lose their protection and may be directly attacked as if they were combatants. Contrary to combatants, however, they do not benefit from immunity from prosecution for lawful acts of war (so-called "combatant privilege") and therefore can be punished by their captor for any violation of national law. Civilians deprived of their liberty—including those having directly participated in hostilities—are entitled to humane treatment and fair trial guarantees as reflected in the various applicable instruments of IHL.[50]

Members of organized armed groups

In IHL governing non-international armed conflict, organized armed groups constitute the armed forces (the armed wing) of a non-state belligerent and must not be confused with the belligerent party itself (for example, an insurgency as a whole, including its political or administrative wing) or with other supportive segments of the civilian population. Treaty IHL governing non-international armed conflict uses the terms civilian, armed forces and organized armed group without defining them. It is generally recognized, however, that members of state armed forces do not qualify as civilians, and the wording and logic of Protocols I and II and of Article 3 common to the Geneva Conventions suggest that the same applies to members of organized armed groups.

Civilians may support a non-state party in various ways and may even participate directly in hostilities on a spontaneous, sporadic or unorganized basis. However, they cannot be regarded as members of an organized armed group unless it is their function to directly participate in hostilities on behalf of the non-state party. Such a combat function does not imply entitlement to combatant privilege, prisoner-of-war status, or any other form of immunity from domestic prosecution for lawful acts of war. Rather, it makes a strictly functional distinction between members of the organized fighting forces and the civilian population. For the present context this means that individuals conducting cyber operations on behalf of a non-state party lose their civilian status and become members of that party's "armed forces" only if their operations are conducted on a continual basis and amount to direct participation in hostilities.[51]

Conclusion

As far as international law is concerned, the phenomenon of cyberwarfare does not exist in a legal vacuum but is subject to well-established rules and principles. However, transposing these rules and principles to the new domain of cyberspace encounters certain difficulties and raises a number of important questions. Some of these questions can be resolved through classic treaty interpretation and common sense; others will require a unanimous policy decision by the international community of states. Cyberwarfare has not yet had dramatic humanitarian consequences—and hopefully this will not change in the future. The potential for human tragedy, however, is enormous, and it is likely to increase with our growing dependence on computer-controlled systems to sustain our daily lives. It is all the more important that states be aware of not only their legal duty to examine whether new weapons and methods employed in cyberwarfare are compatible with their obligations under existing IHL, but also of their moral responsibility towards generations to come.

Notes

1. This article is an excerpt from N. Melzer, *Cyberwarfare and International Law*, UNIDIR, 2011. The original article also examines *jus ad bellum* and the law of neutrality, among other topics.

2. See the International Committee of the Red Cross extensive study on customary IHL, J.-M. Henckaerts and L. Doswald-Beck, *Customary International Humanitarian Law*, vols. I and II, 2005.
3. Shanghai Cooperation Organization, *Annex I to the agreement between the governments of the member states of the Shanghai Cooperation Organization on cooperation in the field of international information security*, 16 June 2009, based on an unofficial translation.
4. M. Schmitt, "Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework", *Columbia Journal of Transnational Law*, vol. 37, 1999, pp. 885–937.
5. Participants at the 2004 Stockholm Expert Conference agreed that "International Humanitarian Law applies to computer network attacks (CNA) in an ongoing international armed conflict". See K. Byström (ed.), *Proceedings of the Conference: International Expert Conference on Computer Network Attacks and the Applicability of International Humanitarian Law, 17–19 November 2004, Stockholm, Sweden*, 2005, p. 181. At the time of writing, the same approach is being taken (unanimously) in the draft *Tallinn Manual*.
6. See Article 22 of the Convention with Respect to the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. Similar phrasing can be found in Article 35 of Protocol I.
7. Security Council, *Report of the Secretary-General pursuant to paragraph 2 of Security Council resolution 808 (1993)*, UN document S/25704, 3 May 1993.
8. ICRC, "How is the Term 'Armed Conflict' Defined in International Humanitarian Law?", *ICRC Opinion Paper*, 2008, p. 5.
9. Ibid., p. 1.
10. See: M. Schmitt, "Cyber Operations and the Jus in Bello: Key Issues", *Naval War College International Law Studies*, vol. 87, 2011, pp. 89–110; and K. Dörmann, "The Applicability of the Additional Protocols to Computer Network Attacks: An ICRC Viewpoint", in K. Byström (ed.), *Proceedings of the Conference: International Expert Conference on Computer Network Attacks and the Applicability of International Humanitarian Law, 17–19 November 2004, Stockholm, Sweden*, 2004, pp. 139–53. At the time of writing, this approach is also taken in the draft Tallinn Manual.
11. K. Dörmann, "The Applicability of the Additional Protocols to Computer Network Attacks: An ICRC Viewpoint", in K. Byström (ed.), *Proceedings of the Conference: International Expert Conference on Computer Network Attacks and the Applicability of International Humanitarian Law, 17–19 November 2004, Stockholm, Sweden*, 2004, p. 142. This approach is also being taken in the draft Tallinn Manual.
12. Article 51(2), Protocol I.
13. Article 52(1), Protocol I.
14. Article 51(4), Protocol I.
15. Article 52(2), Protocol I.
16. Articles 57 and 58, Protocol I, respectively.
17. Article 12(1), Protocol I.
18. Article 41(1), Protocol I.
19. Article 56, Protocol I.
20. Article 44(3), Protocol I.
21. Article 39(2), Protocol I.
22. See the discussion of direct participation in hostilities in relation to collective operations and preparatory measures in N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009, pp. 54–55, 65–67.
23. K. Dörmann, "The Applicability of the Additional Protocols to Computer Network Attacks: An ICRC Viewpoint", in K. Byström (ed.), *Proceedings of the Conference: International Expert Conference on Computer Network Attacks and the Applicability of International Humanitarian Law, 17–19 November 2004, Stockholm, Sweden*, 2004, pp. 139–53.

24. M. Schmitt, "Wired Warfare: Computer Network Attack and *Jus in Bello*", *International Review of the Red Cross*, vol. 84, no. 846, 2002, pp. 365–99.
25. Article 51(5)(b), Protocol I.
26. Article 48, Protocol I.
27. Article 51(1) and (3), Protocol I. Article 13(1) and (3), Protocol II.
28. Article 57(1), Protocol I.
29. Article 37, Protocol I.
30. Article 51(3), Protocol I. Article 13(3), Protocol II.
31. N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009, pp. 47–50.
32. Ibid., pp. 43–44.
33. Article 51(3), Protocol I. Article 13(3), Protocol II.
34. For further information see N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009.
35. Following an ICRC expert meeting, it was agreed that cyber operations directly causing military harm to the adversary in a situation of armed conflict amounted to direct participation in hostilities. See ICRC, *Third Expert Meeting on the Notion of Direct Participation in Hostilities: Summary Report*, 2005, p. 14.
36. Article 48, Protocol I.
37. For a more detailed examination see N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009. For critiques and the ICRC's response see R. Goodman and D. Jinks, "The ICRC Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law: An Introduction to the Forum", *Journal of International Law and Politics*, vol. 42, no. 3, 2010, pp. 637–40.
38. Article 44(3), Protocol I.
39. Article 50(1), Protocol I. See also the broader discussion on the presumption of civilian protection in N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009, pp. 74–76.
40. According to Article 51(4) of Protocol I, indiscriminate attacks are: "(a) those which are not directed at a specific military objective; (b) those which employ a method or means of combat which cannot be directed at a specific military objective; or (c) those which employ a method or means of combat the effects of which cannot be limited as required by this Protocol; and consequently, in each such case, are of a nature to strike military objectives and civilians or civilian objects without distinction".
41. The definition of indiscriminate attack also includes, in Article 51(5)(b) of Protocol I, those "which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated".
42. See the rejection of this view in M. Schmitt, "Cyber Operations and the Jus in Bello: Key Issues", *Naval War College International Law Studies*, vol. 87, 2011, p. 8.
43. Article 49(1), Protocol I. This does not exclude that such operations, as well as the destruction caused in their course, may still amount to an internationally wrongful act.
44. M. Schmitt, "Cyber Operations and the Jus in Bello: Key Issues", *Naval War College International Law Studies*, vol. 87, 2011, p. 89–110; and M. Schmitt, "Wired Warfare: Computer Network Attack and *Jus in Bello*", *International Review of the Red Cross*, vol. 84, no. 846, 2002, pp. 365–99.
45. K. Dörmann, "The Applicability of the Additional Protocols to Computer Network Attacks: An ICRC Viewpoint", in K. Byström (ed.), *Proceedings of the Conference: International Expert Conference on Computer Network Attacks and the Applicability of International Humanitarian Law, 17–19 November 2004, Stockholm, Sweden*, 2004, pp. 139–53.

46. A botnet is an interconnected series of compromised computers used for malicious purposes. A computer becomes a bot when it runs a file that has bot software embedded in it. IP spoofing refers to the creation of internet protocol (IP) packets with a forged source address to conceal the identity of the sender or impersonate another computing system.
47. Taken from Article 37(1), Protocol I.
48. See Articles 38 and 39, Protocol I.
49. Article 37(1), Protocol I.
50. In international armed conflict, civilians deprived of their liberty are protected by the Geneva Convention relative to the Protection of Civilian Persons in Time of War, Protocol I and customary law. In non-international armed conflict these protections are reflected in Article 3 common to the Geneva Conventions, Protocol II and customary law. Depending on the context, human rights law may additionally be relevant.
51. For the ICRC's position on this issue see N. Melzer, *Interpretive Guidance on the Notion of Direct Participation in Hostilities under International Humanitarian Law*, ICRC, 2009.

Cyber offence and defence as mutually exclusive national policy priorities

Brian Weeden

When military strategists consider how best to attack an adversary, the normal course of action is to attack a vulnerability ("weak point") in an opponent's systems. In traditional warfare, identification of potential weak points to exploit would include examination of the adversary's tanks, airplanes, ships, missiles and other types of military hardware for vulnerabilities such as the turning radius of a fighter aircraft or acoustic blind spot of a submarine. Specific tactics are developed to exploit these vulnerabilities, and in some cases specific weapons are built against them. But it is very rare for adversaries to be using the same systems as your own, and thus vulnerabilities in your opponent's systems are not normally also found in your own. Hardening your own systems against vulnerabilities usually does not impact your ability to exploit vulnerabilities in an adversary.

The situation is different in the case in cyberwarfare. The entire cyber domain is built on the foundation of hardware with common processing architectures connected by a standardized system for exchanging packetized data. On top of this, the forces of economics and ease of use have created virtual monocultures in operating systems and popular software. Thus, at the top level, many of the same cyber vulnerabilities exist across many organizations and countries, and this presents a policy dilemma for states when developing and using cyber weapons while also trying to strengthen their cyber defences, both of which have become prominent policy concerns.

Bugs, vulnerabilities and exploits

To see why this dilemma exists, we first have to take a step back and understand how cyber weapons are developed. Programming errors, colloquially known as "bugs", are a fact of life when creating software. A bug is commonly defined as the difference between what a programmer meant for a piece of code to do and what the code actually does. This difference can come about via a number of different ways, some of which are extremely difficult to detect. It is difficult and expensive to write a bug-free software program, and the difficulty approaches impossibility as software becomes more complex, such as many modern programs that contain millions of lines of codes developed by hundreds or thousands of programmers. Many bugs will simply cause programs to crash, but sometimes they can be used to perform an action that is normally not allowed. These are known as vulnerabilities, and the tactic for making use of vulnerabilities to compromise a computer or to access information is known as an exploit. It is important to note that vulnerabilities can also arise through flaws in the design of a system, such as a purposeful shortcut taken for ease of use or improper implementation of an encryption algorithm.

Brian Weeden is a former US Air Force officer with experience in space control and nuclear operations. He is currently a PhD student in Science and Technology Policy at George Washington University.

Although the end target of a cyber attack may be a very specific system or piece of equipment, it is almost certainly attached to a more common system such as a desktop computer or server that is running one of the ubiquitous operating systems or software packages. In many cases these host computers are connected via a local network to other computers. Just as with developing any other weapon, developing a cyber weapon begins by determining what the specific target is and the effect desired by the attack. The specific hardware and software in the computers connected to the target are studied for any potential vulnerabilities and when found one or more exploits for the vulnerabilities are developed. The cyber weapon contains software and routines that use exploits to attack and infiltrate each level of the system from injection point to the target. Once the final target system is reached, an exploit crafted specifically to do some form of harm or damage to that system, commonly known as a payload, is deployed.

Thus, an important step in developing cyber weapons involves discovering new vulnerabilities in software or hardware and developing related exploits. Many exploits will utilize remote code execution (the ability for an attacker to execute code of their choosing on the target system) or privilege escalation (the ability for an attacker to obtain higher privileges on the system to execute commands). Most prized of all are so-called "zero-day" exploits, which are so named because the first time they are publicly known to exist is when they are found being used "in the wild" as part of a cyber attack. Militaries, intelligence agencies, organized crime and "black hat" hackers are all working constantly to find and compile libraries of exploits that can be used for cyber offence. At the same time, software vendors, security researchers and other "white hats" are racing to find vulnerabilities so that they can be patched to improve cyber defence.

Although all software can have vulnerabilities, the efforts of both black hats (hackers that have ostensibly malicious intentions) and white hats (hackers that are working to improve security) are often focused on the most popular applications so that the largest possible number of hosts can be targeted (or protected). At approximately 85% market share of operating systems, Microsoft Windows presents a very lucrative target.[1] Likewise, the very high market share for Microsoft Office, especially in corporate and government environments, makes it an important target as well. Even more compelling are applications that exist across all platforms, such as Sun Microsystem's Java platform and Adobe's PDF Reader and Flash software.[2]

Case example: Stuxnet

The Stuxnet malware, used in attacks on the Iranian nuclear enrichment facility at Natanz, is a good example of how a cyber weapon is developed and deployed, and creates the choice between cyber offence and defence. Originally detected by security researchers in June 2009, the Stuxnet attacks consisted of multiple versions of a complex Microsoft Windows malware discovered up until mid-2010.[3] Although the definition of what constitutes a cyber weapon is

slippery at best, many cybersecurity experts consider the Stuxnet attacks to represent the first public demonstration of what a real-world cyber weapon can do. [4]

The ultimate targets for Stuxnet were the centrifuges used at Natanz for enriching uranium. To affect the centrifuges, Stuxnet needed to infect computers known as industrial control systems (ICS) used to program and control the programmable logic controller devices (PLCs) which in turn controlled the frequency converter drives that ran the centrifuges.[5] This meant that Stuxnet first needed to exploit vulnerabilities in the host operating system to infect the overall machine, exploit vulnerabilities in the application software for the PLCs, exploit vulnerabilities in the PLC themselves, and finally command the frequency converters in a way that damaged the centrifuges.[6]

The ICS computers Stuxnet needed to infect were not connected directly to the Internet—they used a common security protocol called "air gapping" (physical, electrical and electromagnetic isolation) to insulate them from other systems and in particular the Internet. However, operators still needed a way to update the software on these computers and transfer data to and from them. As is often the case, this was done using removable USB thumb drives. Stuxnet was designed to exploit this practice by infecting USB thumb drives and spreading peer-to-peer between computers within a local network.[7] Stuxnet was unleashed in three different waves against five different organizations with a presence in Iran.[8] Over a period of time, Stuxnet spread within and between networks until finally reaching the ICS computers, where the payload executed.

All of the computers that Stuxnet infected were running Microsoft Windows, and Stuxnet took advantage of four zero-day exploits in Windows to infiltrate its targets. The first version of Stuxnet, discovered in June 2009, took advantage of a remote code execution vulnerability in the Windows Print Spooler Service.[9] This vulnerability had been previously disclosed by the security magazine *Hakin9* in April 2009, but was not patched by Microsoft until September 2010.[10] A new version of Stuxnet, discovered in March 2010, exploited a previously unknown remote code execution vulnerability in the way Windows handles shortcut or link files.[11] Microsoft issued a security advisory for this vulnerability in July 2010 and a patch to fix it in August 2010.[12] The security firm Symantec privately disclosed two other privilege escalation vulnerabilities to Microsoft as a result of Symantec's analysis of Stuxnet.[13]

Stuxnet's developers either discovered these vulnerabilities in Windows and the other parts of the system in the process of developing Stuxnet or had a library of publicly unknown exploits on hand to draw from. In either case, knowledge of these vulnerabilities was kept secret and not disclosed to Microsoft. This left many millions of computers owned by governments, companies and private citizens around the world vulnerable to the same exploits, while the eventual deployment and discovery of Stuxnet made the code for these exploits publicly available.

The developers of Stuxnet did take considerable steps to limit its spread. Unlike many other types of malware, Stuxnet was not a worm—it was not designed to spread over the open internet as fast or as widely as possible. Stuxnet only spread via USB thumb drives and within a local area network (LAN), and each infected device was limited to infecting three others.[14] Stuxnet also contains code that suggests it will "self-destruct" on 24 June 2012, although there is debate among security experts on whether this code will actually work or what it will do.[15]

Despite these constraints, as of September 2010 Stuxnet had infected over 100,000 hosts in 155 countries.[16] Although the majority of these infections were in Iran, significant infections were also found in Indonesia and India. This spread occurred because of the still unpatched vulnerabilities in Windows and the widespread use of unsafe practices regarding USB thumb drives. In October 2010, Siemens confirmed that 15 of its industrial customers—including chemical plants, power plants and production facilities—located around the world had been "affected" by Stuxnet.[17] These customers used Siemens PLCs to control various systems, although it is unknown if Stuxnet caused any damage.

To be clear, although Stuxnet infected all of these systems there is no evidence that it did any harm or damage to any systems outside of Iran. Analysis of the Stuxnet code performed by the security community has shown that its malicious payload was crafted to only execute against the specific ICS computers used for the Iranian centrifuges at Natanz. The Windows machine that it was targeted to infect needed to be running the Step 7 software used to control PLCs manufactured by Siemens Corporation. The PLCs needed to be a Siemens model 6ES7-315-2 controlling at least 33 frequency converter drives, manufactured by Fararo Paya in Tehran or by Vacon in Finland, running between 807 and 1,210 Hz.[18]

However, once released into the wild, the inevitable spread of Stuxnet made it possible for anyone with the tools and motivation to discover how it worked, and the time lag of months to years before the vulnerabilities it used were patched allowed plenty of time for organized crime and other black hats to take advantage. At the end of July 2011, Microsoft reported a massive spike in the number of malware infection attempts using the same shortcut/link exploit used by Stuxnet.[19] These attempts occurred all over the world but were especially prevalent in Brazil and the United States, which previously did not have a large number of Stuxnet infections. Currently, there is another malware in the wild, known as Duqu, which bears a striking resemblance to Stuxnet, leading some security researchers believe it is from the same developers or was built by re-using key parts of Stuxnet.[20]

National cyber policy choices for states

As the Stuxnet example shows, governments that are seeking both to strengthen their own national cyber defences and to develop offensive cyber techniques and weapons that can be used against adversaries are faced with a conundrum arising from the unique nature of the cyber domain. Pursuing the usual offence–defence chess match in regards to cybersecurity

would have impacts outside of the military domain because of the widespread use of the same software and hardware across military, commercial and civilian applications. Furthermore, it is very likely that any useful exploits that militaries or intelligence agencies discover in developing offensive cyber weapons could also be used against their own systems as well as those of commercial companies and private citizens, leading to a policy choice of either favouring cyber offence or cyber defence. The following sections examine some of the issues faced in promoting offence or promoting defence.

Favouring cyber offence

From the offensive perspective, the traditional policy choice would be to classify any vulnerabilities (in hardware, software, or systems) to enable development of exploits and eventually offensive cyber capabilities and also to keep them out of the hands of other states and cybercriminals. This is the traditional choice that governments make when it comes to conducting offensive military campaigns in any domain. Although it can be successfully accomplished in traditional domains of warfare, in the cyber domain this choice falsely assumes that only governments are involved in defence and offence. That is a reasonable assumption in the case of land, sea, or air operations, but not in the cyber domain.

Part of the difficulty in cyber defence is the extent of the "attack surface" that must be defended. Vulnerabilities exist in many more places than just operating systems and in many more objects than just traditional desktop and laptop computers. Virtually everything that runs computer software is likely to have a vulnerability, especially categories of software that have not been targeted in the past, such as PLCs or, more recently, mobile phone operating systems. There are several reported cases of researchers finding a specific vulnerability in a software package whenever they go looking for one.[21] After Stuxnet, PLCs and other gear used to control industrial processes have received increased attention. In one case, a security researcher discovered several new critical ICS vulnerabilities, which he planned to unveil at a cybersecurity conference. However, the vulnerabilities he discovered were serious enough that he was persuaded by the US Department of Homeland Security and Siemens to forego his talk.[22] And as more and more devices become "smart", the number of potential vulnerabilities is growing at an increasing rate. For example, at the 2011 Black Hat Conference (one of the most popular and notable hacker meetings), researchers demonstrated a technique for remotely unlocking and starting a car using text messages.[23]

It is also likely that offensive cyber operations would be classified as covert operations—activities that are planned and executed so as to conceal the identity of, or permit plausible denial by, the sponsor or sponsors. This often means knowledge about the cyber operation would be compartmentalized from other organizations even within the same government, furthering the likelihood that its own systems would be vulnerable to the same exploits in a manner similar to the following example.

For a variety of logistical, organizational and security reasons the US military operates dozens of computer networks. Members of the armed forces often need to access and transfer data across multiple networks to perform their mission, a need complicated by the fact that many of these systems are air gapped from each other for security. Using removable drives to transfer data between various networks, and in many cases bypassing security features and protocols, is thus an operational necessity.

The US military banned the use of all removable drives on the unclassified NIPRNET and classified SIPRNET networks in November 2008. The ban was a response to a significant infection of the Agent.btz malware in those networks stemming from a single USB thumb drive that was picked up in a parking lot of a base in the Middle East.[24] The malware in question was a worm that infected and spread through various versions of Microsoft Windows. The US military reinstated the ability to use removable drives under specific situations in February 2010, after an extensive cleaning operation called BUCKSHOT YANKEE.[25] Shortly thereafter, Bradley Manning allegedly used removable media in the form of writable compact discs to remove 150,000 diplomatic cables from a secure facility in Iraq, cables that would later be published by Wikileaks for the world to see. In December 2010, the US military once again re-instated the ban on removable drives.

Although there are no known links between Stuxnet and Agent.btz, the similarities of the attack profile are striking. There is widespread speculation and significant circumstantial evidence that the United States was either behind or complicit in Stuxnet,[26] which was likely being developed during the time period that Agent.btz was infecting the US military networks. If elements of the US government were involved in the development of Stuxnet, it would appear that there was no communication between the entities developing Stuxnet for offence and those that were responsible for defending US military networks as to the extreme vulnerability posed by using removable drives on machines running Microsoft Windows and the ability for malware using removable drives to bypass air gaps. Although such a lack of communication and coordination among US government agencies would be similar to the intelligence failures prior to the 11 September 2001 attacks,[27] in the cyber domain such a lack of communication or coordination is even more likely because of the incentive to prevent a compromise of offensive capabilities or operations.

Thus, policies aimed at improving a state's cyber offensive capabilities would hinder the ability of that state to improve its cyber defence and result in counter-productive efforts, bureaucratic infighting and significant duplication of resources. There would be constant trade-offs between revealing a vulnerability to vendors, industry and the public so that they can be fixed, and keeping the vulnerability classified so that it can be potentially used offensively by the military, intelligence agencies or law enforcement. Both US and German law enforcement have developed and deployed software for surveillance purposes that use vulnerabilities and exploit techniques similar to those used by cybercriminals,[28] and an entire industry has cropped up around providing vulnerabilities to governments rather than to

vendors or the public.[29] At an annual hacking contest sponsored by Google in March 2012, a well-known cybersecurity firm refused to divulge vulnerabilities and exploits that it had discovered in Google's Chrome web browser because it was worth more to them to divulge such information only to their customers, which consist of NATO governments and NATO partners.[30] Such behaviour has led to debates among cybersecurity firms as to whether they should warn end-users about government-designed malware.[31] Doing so might compromise ongoing investigations and place the firms in legal jeopardy with governments, but not doing so allows other governments and cybercriminals to use the same techniques unnoticed.

Favouring cyber defence

From the defensive perspective, a logical policy option would be to develop and acquire custom software for use in critical national security or infrastructure applications. While this option would force an attacker to discover vulnerabilities in those systems, it has some significant drawbacks. No piece of software of any significant size is ever secure "out of the box". All software ships with a number of bugs, of which a small percentage could be security vulnerabilities. Simply reviewing the code is not good enough. Even in the case of open source software where theoretically anyone can view the source code and find bugs, there have been instances of significant bugs in security protocols going undiscovered for many years.[32] The only way to find and correct bugs is to test the software under all possible conditions of use. This is an extraordinarily time consuming undertaking for complex software programs, and in some cases can be economically impractical or even impossible in practice.

Using a popular piece of software has an advantage in that it gets much more use under a lot of different circumstances, and is also likely being attacked more often. This leads to faster discovery and elimination of bugs and thus of bug-based vulnerabilities. Evidence of this can be seen in the decrease in the number of significant bug-based vulnerabilities in popular software such as web browsers.[33] At the same time, both software manufacturers and users develop a set of security practices which is an important part of cyber defence. A piece of software that is only used by a small number of individuals will still have bugs and vulnerabilities, but it will take much longer for them to be discovered and the developer and users will likely not have developed optimized procedures for dealing with attacks.

A good example of these two situations are the operating systems developed by Microsoft and Apple—Windows and OS X, respectively. Although Microsoft's Windows enjoys a huge market share, it has a long and chequered history when it comes to security. Poor architecture choices and the need to support legacy devices and third party applications resulted in all versions of Windows having a significant number of vulnerabilities. However, as a result Microsoft has developed an excellent process for identifying vulnerabilities, warning users, and developing and rolling out patches. Over time, Windows itself has become much more secure, and its users have also become rather "street wise" and wary of some of the more basic attacks.

Although Apple has seen rapid growth in desktop and laptop market share over the last several years, it still has only a fraction of Microsoft's market share.[34] This small install base meant that attackers initially had less of an economic incentive to exploit vulnerabilities in it. Combined with Apple's aggressive marketing campaign of OS X being a more "secure" operating system than Windows, this has led many of Apple's users to believe that they were in fact safe. The "Mac Defender" attacks first discovered in May 2011, which sent Apple scrambling to respond with a patch and new policies, dispelled any notions of Apple having a more secure platform. Apple is now locked in the same cat-and-mouse war with attackers as Microsoft, albeit without the latter's years of experience in doing so and with a user base that believes Apple devices are inherently safe.[35]

It is possible to take existing off-the-shelf commercial or open source software packages and harden them through code auditing and minimizing software functionality by eliminating unneeded functions and capabilities.[36] For a defence-focused cyber policy, this is likely the best option to protect critical networks and capabilities—if done in an intelligent manner, it can drastically improve security for much lower costs than developing custom software and still leverage the robustness of popular software.[37] However, no system can ever be proven completely secure, and widespread use of this tactic will simply invite closer examination of other entry points on the attack surface. Most notably, the humans in a system are frequently the most vulnerable piece, as shown by the high success rate of social engineering attacks, such as phishing, utilizing malicious email attachments. In addition to the Agent.btz infection of US military networks previously mentioned, recent major attacks on security firms RSA and HBGary,[38] Google and other technology companies,[39] and dozens of other government and private entities[40] all exploited the humans in the system.

A significant element of cyber defence is increased cooperation and coordination among governments, private industry and academia. Commercial software applications and architectures are widely used throughout government computer systems and networks, making the protection of government systems in part reliant on discovering and fixing vulnerabilities in commercial software. Additionally, the private sector itself represents a significant part of the attack surface a state needs to defend to protect against cyber attacks, or at least to mitigate the consequences. Private companies operate significant parts of critical national infrastructure, all of which are potential targets of cyber attack.

Academics already play a significant role in the cyber security world, but efforts by researchers are often hindered or stigmatized by corporations and governments because they are seen as a threat and not an asset.[41] This would need to change, and the cyber community would need to adopt a favourable attitude towards any research and experimentation that leads to a better understanding of cyber vulnerabilities and weaknesses in security architectures.

The public is often overlooked but plays a potentially significant role in cyber defence. The many millions of personal computers are potential weapons that can be compromised by an

attacker and turned into weapons, for example as part of a botnet running a denial of service attack. Compromised personal computers, mobile devices or online accounts of government officials and corporate executives could provide critical information that leads to the compromise of protected systems. Friends and relatives on social networks are also potential avenues of attack, potentially more likely to succeed because of their trusted nature.

Thus, policies aimed at improving a state's cyber defence would necessarily need to increase the amount of information-sharing among governments, industry, academia and potentially even the public, and make major changes in the current classification policy for cyber vulnerabilities and attacks. Governments, industry and academia would need to share information about the latest attacks, malware signatures and vulnerabilities.[42] Incentive programmes for the responsible disclosure of vulnerabilities, such as those already being run by Google and the Mozilla Foundation for their respective web browsers,[43] could greatly increase the number of people looking for vulnerabilities and the rate at which they are discovered and fixed. However, these approaches would also have an increasingly negative impact on the ability of a state to develop and field offensive cyber capabilities over time, largely through the increased cost of finding new vulnerabilities and developing offensive weapons against them even as they are being patched.

Conclusions

Cybersecurity presents unique challenges to policymakers, especially when it comes to dealing with the twin goals of protecting one's own networks while simultaneously developing tools and techniques to attack the networks of adversaries. Stuxnet used four previously unknown vulnerabilities in Microsoft Windows to infect its targets, vulnerabilities which were present in hundreds of millions of computers around the world and, once disclosed, were open for exploitation. The developers of Stuxnet chose to use these vulnerabilities for offence, instead of disclosing them to security firms and software vendors so they could be fixed, enabling other cyber actors to exploit the same vulnerabilities across a range of malware and attacks against governments, companies and citizens. While Microsoft has since patched all four Windows vulnerabilities exploited by Stuxnet, some of the major design vulnerabilities in the Siemens ICS code exploited by Stuxnet have not been fixed as of January 2012.[44]

The example of Stuxnet demonstrates the difficult choices national policymakers must make if they wish to pursue cyber offence, as doing so means giving cyber defence a lower priority. Trying to keep vulnerabilities secret so that they can be used for offence will likely result in vulnerabilities going unfixed, thus hampering defensive efforts. Developing custom software based on open source or commercial software for use in critical national security and infrastructure applications in some ways breaks the link between offence and defence, but doing so will incur significant costs. More importantly, it would provide protection for only those systems running such custom software—the systems of allies, commercial companies and citizens would fall outside its protection. Likewise, policies that favour cyber defence (such

as bounty programmes for finding vulnerabilities) and increased cooperation, coordination and information-sharing among governments, industry, academia and the public would make it more difficult for a state to develop and maintain classified or covert offensive cyber programmes.

Ultimately, states must make a choice between prioritizing cyber offence or cyber defence. Both cannot be done well at the same time, and focusing on one lessens the ability to successfully accomplish the other. Although there is an increasingly loud cry for increased cybersecurity from virtually all states, both stated policy and unstated actions make it clear that many states are currently giving priority to cyber offence—in particular the US government has announced a number of initiatives to speed up military development of offensive cyber weapons.[45] This will likely need to change if states want to match their rhetoric on the need for cyber defence with meaningful actions to protect themselves and their citizens.

Notes

1. "Operating System Market Share", NetMarketShare, March 2012, <http://marketshare.hitslink.com/operating-system-market-share.aspx?qprid=8>.
2. *Cisco 2010 Annual Security Report*, Cisco Systems, 2011, <www.cisco.com/en/US/prod/collateral/vpndevc/security_annual_report_2010.pdf>, p. 22.
3. For detailed information, see "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>.
4. See, for example, Gary D. Brown, "Why Iran Didn't Admit Stuxnet Was an Attack", *Joint Force Quarterly*, no. 63, 2011, <www.ndu.edu/press/why-iran-didnt-admit-stuxnet.html>; and Sydney J. Freedberg Jr., "Cyber Command Lawyer Praises Stuxnet, Disses Chinese Cyber Stance", AolDefense, 12 March 2012, <http://defense.aol.com/2012/03/12/cyber-command-lawyer-praises-stuxnet-disses-chinese-cyber-stanc/>.
5. See "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>.
6. See "Enumerating Stuxnet's exploits", Langner Communications, 7 June 2011, <www.langner.com/en/2011/06/07/enumerating-stuxnet%E2%80%99s-exploits/>.
7. "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>, p. 2.
8. Ibid., p. 9.
9. Ibid., p. 4.
10. See "Microsoft Security Bulletin MS10-061 - Critical", Microsoft, 14 September 2010, <www.microsoft.com/technet/security/Bulletin/MS10-061.mspx>.
11. "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>, p. 4.
12. See "Microsoft Security Bulletin MS10-046 - Critical", Microsoft, 2 August 2010, <www.microsoft.com/technet/security/bulletin/MS10-046.mspx>.
13. "Updated W32.Stuxnet Dossier is Available", Symantec, updated 14 February 2011, <www.symantec.com/connect/blogs/updated-w32stuxnet-dossier-available>.
14. "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>, p. 10.

15. Ibid., p. 18; and Michael Joseph Gross, "A Declaration of Cyber-War", *Vanity Fair*, April 2011, <www.vanityfair.com/culture/features/2011/04/stuxnet-201104>.
16. "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>, p. 5
17. "Cyber worm found at German industrial plants", *The Local*, 2 October 2010, <www.thelocal.de/national/20101002-30225.html>.
18. See "W32.Stuxnet Dossier", Symantec, ver. 1.4, 2011, <http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf>; and Dale G. Peterson, "Langner's Stuxnet Deep Dive S4 Video", Digital Bond, 31 January 2012, <www.digitalbond.com/2012/01/31/langners-stuxnet-deep-dive-s4-video/>.
19. Holly Stewart, "Stuxnet, malicious .LNKs, ... and then there was Sality", Microsoft Makware Protection Center, 30 July 2010, <http://blogs.technet.com/b/mmpc/archive/2010/07/30/stuxnet-malicious-lnks-and-then-there-was-sality.aspx>.
20. See "W32.Duqu", Symantec, ver. 1.4, 23 November 2011, <www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_duqu_the_precursor_to_the_next_stuxnet.pdf>.
21. See, for example, "This is how Windows get infected with malware", CSIS Security Group, 27 September 2011, <www.csis.dk/en/csis/news/3321/>; and Kim Zetter, "Researchers release new exploits to hijack critical infrastructure", Ars Technica, 5 April 2012, <http://arstechnica.com/business/news/2012/04/researchers-release-new-exploits-to-hijack-critical-infrastructure.ars>.
22. Chris Blask, "Network Security: The Threats You Don't See", Infosec Island, 22 June 2011, <www.infosecisland.com/blogview/14682-Network-Security-The-Threats-You-Dont-See.html>.
23. See Don Andrew Bailey, "War Texting", <www.isecpartners.com/storage/docs/presentations/iSEC_BH2011_War_Texting.pdf>.
24. See William J. Lynn III, "Defending a New Domain", *Foreign Affairs*, vol. 89, no. 5, 2010, <www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain>.
25. See Ellen Nakashima, "Cyber-intruder sparks massive federal response — and debate over dealing with threats", *Washington Post*, 9 December 2011, <www.washingtonpost.com/national/national-security/cyber-intruder-sparks-response-debate/2011/12/06/gIQAxLuFgO_story.html>; and Leo Shane III, "DOD loosens restrictions on thumb drives", *Stars and Stripes*, 19 February 2010.
26. See, for example, Tom Gjelten, "Security Expert: U.S. 'Leading Force' Behind Stuxnet", National Public Radio, 26 September 2011, <www.npr.org/2011/09/26/140789306/security-expert-u-s-leading-force-behind-stuxnet>; Ron Rosenbaum, "Richard Clarke on Who Was Behind the Stuxnet Attack", *Smithsonian*, April 2012, <www.smithsonianmag.com/history-archaeology/Richard-Clarke-on-Who-Was-Behind-the-Stuxnet-Attack.html>; Michael Joseph Gross, "A Declaration of Cyber-War", *Vanity Fair*, April 2011, <www.vanityfair.com/culture/features/2011/04/stuxnet-201104>; and Dale G. Peterson, "Langner's Stuxnet Deep Dive S4 Video", Digital Bond, 31 January 2012, <www.digitalbond.com/2012/01/31/langners-stuxnet-deep-dive-s4-video/>.
27. *The 9/11 Commission Report*, 2004, p. xvi and passim.
28. Kevin Poulsen, "FBI Spyware: How Does the CIPAV Work? — UPDATE", Wired, 18 July 2007, <www.wired.com/threatlevel/2007/07/fbi-spyware-how/>; and Matthew Lasar, "Impressed by FBI trojan, Germans write their own—and national scandal ensues", Ars Technica, 14 October 2011, <http://arstechnica.com/security/news/2011/10/impressed-by-fbi-trojan-germans-write-their-ownand-national-scandal-ensues.ars>.
29. Michael Ray and Ashlee Vance, "Cyber Weapons: The New Arms Race", *Businessweek*, 20 July 2011, <www.businessweek.com/printer/magazine/cyber-weapons-the-new-arms-race-07212011.html>
30. Andy Greenberg, "Meet The Hackers Who Sell Spies The Tools To Crack Your PC (And Get Paid Six-Figure Fees)", *Forbes*, 21 March 2012, <www.forbes.com/sites/andygreenberg/2012/03/21/meet-the-hackers-who-sell-spies-the-tools-to-crack-your-pc-and-get-paid-six-figure-fees/>.
31. John Leyden, "AV vendors split over FBI Trojan snoops", The Register, 27 November 2001, <www.theregister.co.uk/2001/11/27/av_vendors_split_over_fbi/>; and Stewart Mitchell, "F-Secure: security firms should block

state malware", PC Pro, 8 march 2011, <www.pcpro.co.uk/news/security/365791/f-secure-security-firms-should-block-state-malware>.

32. See, for example, Jake Edge, "A hole in crypt_blowfish", LWN.net, 22 June 2011, <http://lwn.net/Articles/448699/>.
33. Robert Lemos, "The End Of Vulnerabilities?", Dark Reading, 15 March 2012, <www.darkreading.com/vulnerability-management/167901026/security/security-management/232602714/the-end-of-vulnerabilities.html>.
34. See "Operating System Market Share", NetMarketShare, March 2012, <http://marketshare.hitslink.com/operating-system-market-share.aspx?qprid=8>.
35. Don Reisinger, "Mac OS X Security Must Become a Priority: 10 Reasons Why", eWeek, 5 April 2012, <www.eweek.com/c/a/Security/Mac-OS-X-Security-Must-Become-a-Priority-10-Reasons-Why-705108>.
36. Leander J Brandt IV, "Defending the Cyber Alamo: An Indefensible Position in Cyberspace", *High Frontier*, vol. 7, no. 3, 2011, <www.afspc.af.mil/shared/media/document/AFD-110519-023.pdf>, p. 24.
37. Thor Olavsrud, "Do Insecure Open Source Components Threaten Your Apps?", NetworkWorld, 30 March 2012, <www.networkworld.com/news/2012/033012-do-insecure-open-source-components-257846.html>.
38. Uri Rivner, "Anatomy of an Attack", RSA, 1 April 2011, <http://blogs.rsa.com/rivner/anatomy-of-an-attack/>; and Peter Bright, "Anonymous speaks: the inside story of the HBGary hack", Ars Technica, 15 February 2011, <http://arstechnica.com/tech-policy/news/2011/02/anonymous-speaks-the-inside-story-of-the-hbgary-hack.ars>.
39. "Protecting Your Critical Assets", McAfee Labs and McAfee Foundstone Professional Services, 2010, <www.mcafee.com/us/resources/white-papers/wp-protecting-critical-assets.pdf>.
40. Hon Lau, "The Truth Behind the Shady RAT", Symantec, 4 August 2011, <www.symantec.com/connect/blogs/truth-behind-shady-rat>.
41. See, for example, Jaikumar Vijayan, "Carrier IQ drops legal threat against security researcher", Computerworld, 28 November 2011, <www.computerworld.com/s/article/9222203/Carrier_IQ_drops_legal_threat_against_security_researcher>.
42. Jason Healey, "Cybersecurity Legislation Should Force U.S. Government to Listen Less and Speak More", *The Atlantic*, 15 March 2012, <www.theatlantic.com/technology/archive/2012/03/cybersecurity-legislation-should-force-us-government-to-listen-less-and-speak-more/254491/>.
43. "Encouraging More Chromium Security Research", The Chromium Blog, 28 January 2010, <http://blog.chromium.org/2010/01/encouraging-more-chromium-security.html>; and "Bug Bounty Program", Mozilla, 1 February 2012, <www.mozilla.org/security/bug-bounty.html>.
44. See Dale G. Peterson, "Langner's Stuxnet Deep Dive S4 Video", Digital Bond, 31 January 2012, <www.digitalbond.com/2012/01/31/langners-stuxnet-deep-dive-s4-video/>, starting at 45:55.
45. Jim Wolf, "U.S. says will boost its cyber arsenal", Reuters, 7 November 2011, <www.reuters.com/article/2011/11/07/us-cyber-usa-offensive-idUSTRE7A640520111107>; Ellen Nakashma, "U.S. accelerating cyberweapon research", *Washington Post*, 19 March 2012, <www.washingtonpost.com/world/national-security/us-accelerating-cyberweapon-research/2012/03/13/gIQAMRGVLS_story.html>; and Statement of General Keith B. Alexander, Commander, United States Cyber Command before the US House Committee on Armed Services, Subcommittee on Emerging Threats and Capabilities, 20 March 2012, <http://armedservices.house.gov/index.cfm/files/serve?File_id=69276bbe-070a-4b82-8d85-9440931bc8e0>.

Transparency and confidence-building measures in cyberspace: towards norms of behaviour

Ben Baseley-Walker

In 2012 the world's pulse beats in cyberspace. From commerce to development to fighting wars, cyberspace usage is a defining characteristic of our age. Since 2000, with the dramatic increase in the use of cyber technologies in civil, military and commercial sectors, a new, highly dynamic security stage has arisen. Governments are struggling to contend with the security implications of this emerging arena of potential conflict. Today, as an understanding of global dependence on cyber resources has begun to emerge, governments are now taking strong stances on building predictability, stability and security in cyberspace. As can be seen from Stuxnet to attacks on the New York Stock Exchange, cyberspace is now a domain, like sea, air and outer space before it, where fundamental state interests are starting to be expressed. This is a world where terrestrial borders can no longer be said to be the boundaries they once were. Cyberspace has become a new conduit for governmental, as well as non-governmental, power projection.[1]

Following the cyber attacks in Estonia and Georgia in 2007 and 2008 respectively and the attack on Iranian nuclear facilities in 2010, it is becoming increasingly clear that the potential for cyberwarfare has become an "unavoidable element in any discussion of international security".[2] So far at least 33 states now include cyberwarfare in their military planning and organization.[3] There is a growing realization, however, seen in simulations and through political and military analyses, that currently there is little to no ability to effectively control the escalation of cyberconflict. Nor is there any common understanding of how the existing norms of international humanitarian law would apply—if at all.

This article examines a key step on the road to changing that state of affairs—the creation of norms of behaviour and transparency and confidence-building measures (TCBMs). It first examines the nature of TCBMs for cyberspace, their application and then continues with a look at some of the initiatives already proposed.

For the purposes of definition a clear line should be drawn between cybercrime and cyberwarfare. However, the realities of defining the boundaries between different negative activities in cyberspace are complex. Cybercrime can be defined as non-state sponsored actions which are illegal at either the national or international level. This can range from credit card fraud to child pornography. This article, however, deals specifically with cyberwarfare, which is defined as state-sponsored, offensive cyber activities directed towards another state, its infrastructure or population. It is important to note that a common understanding of the

---

Ben Baseley-Walker is Programme Lead of the Emerging Security Threats Programme at the United Nations Institute for Disarmament Research (UNIDIR). He was previously Advisor on Security Policy and International Law for the Secure World Foundation (SWF). The opinions expressed in this article are the author's own and do not necessarily represent the views of UNIDIR or the United Nations.

parameters of the grey area between espionage—illegal data collection—and cyberwarfare has not yet been developed by the international community.

TCBMs: the concept

TCBMs can be broadly defined as elements of international policy that reduce threats, build trust, and make relationships between states more predictable. TCBMs have a long history as a useful tool for the international community and have been used in a variety of international security issues, most notably dealing with nuclear weapons.[4] TCBMs have traditionally been viewed as instruments with a politically binding effect. Although they are usually seen as a bridge to future legally binding international security instruments, the possibility is not precluded that they could themselves become legally binding.

The concept of TCBMs and norms of behaviour has been the subject of much political debate. The terms confidence, security and transparency have been used in various ways, with each concept invariably generating a negative reaction from one state or another. However, the international community has consistently agreed that cyberspace measures of some sort must be taken and taken soon.

The United Nations has long promoted TCBMs as a mechanism to promote security among Member States. In the early 1980s the UN Disarmament Commission developed a set of guidelines for confidence-building measures, which it presented at a special session of the General Assembly devoted to disarmament:

> 2.2.5 A major objective is to reduce or even eliminate the causes of mistrust, fear, misunderstanding and miscalculation with regard to relevant military activities and intentions of other States, factors which may generate the perception of an impaired security and provide justification for the continuation of the global and regional arms build-up.
>
> 2.2.6 A centrally important task of confidence-building measures is to reduce the dangers of misunderstanding or miscalculation of military activities, to help to prevent military confrontation as well as covert preparations for the commencement of a war, to reduce the risk of surprise attacks and of the outbreak of war by accident; and thereby, finally, to give effect and concrete expression to the solemn pledge of all nations to refrain from the threat or use of force in all its forms and to enhance security and stability.[5]

For the purposes of this article TCBMs are measures designed to lessen the likelihood of conflict escalating through a lack of understanding and trust in the cyber activities of both allies and adversaries. While there are many advantages to why a TCBM should be a legally binding measure, given the general state of uncertainty and mistrust between states on cybersecurity issues, it seems likely that TCBMs in cyberspace will be at most only politically binding.

TCBMs generally come in two types: those dealing with capacity and those dealing with intentions. Some states have framed the first in terms of a "duty of care obligation"—a demonstration of best security practices at the state level. The second focuses on international norms and building a better understanding of state-to-state interaction on cyber-related international security issues.[6] Historically, TCBMs have either been constructed to supplement legally binding instruments or have contributed to laying down the foundations for future progress. This can take the form as either progression towards a legally binding instrument or simply an improved climate for building understanding while continuing, for example, to conduct activities in cyberspace or develop cyber defences, and ensuring doctrine on such developments is made widely available.

It is important to emphasize that TCBMs do not necessarily have to be of a particular form or structure. Activities carried out by completely commercial entities, such as the sharing of data on cyber attacks, can amount to a TCBM that clearly fulfils the role of decreasing political and military tensions at the state level. There is a variety of such profit-driven cooperation in other sectors—in the space sector, for example, the sharing of orbital positioning data among commercial satellite operators through the Space Data Association has had a positive impact on the sharing of data and information among government entities.

What do we want to achieve?

It is clear that the goal is to develop a safe, stable and—above all—predictable environment in cyberspace. A state's incentive to inflame tensions or damage the overall cyber environment is inversely proportional to its national engagement in cyberspace. As a state increases its investment in cyber resources—civilian, commercial and military—and derives ever-increasing economic benefit from the Internet, the asymmetrical advantage of attacking an adversary with a heavy reliance on cyber resources risks engendering significant consequences for the perpetrator.

One of the greatest challenges in cyberspace is attribution of an attack. Even if the perpetrator is identified in a timely fashion with a high degree of confidence, proving an act was state-sponsored is extremely challenging and often impossible. This leaves states with few options—either do nothing or risk a crisis situation which might quickly escalate, given that neither side has a clear idea of the "red lines" of their adversary nor a clear understanding of what the escalatory steps might be. This reality means that building established mechanisms of state interaction in cyberspace is essential if we hope to slow escalation and reduce the likelihood of conflict. TCBMs play a central role in reducing misperceptions and communicating the long-term intentions of states.

Understanding the position of allies and adversaries: a first step

One of the first steps to building an effective TCBM regime is to develop a clear understanding of the parameters within which other actors in cyberspace operate. In the political–military realm the development and sharing of military doctrine, an appreciation of the exact aims of a national declaratory policy on cyberspace and creating crisis management links, such as hotlines, are all essential. Certain political issues are shaping up to be highly contentious. The most current of these is the question of whether information can be viewed as a weapon. Some states view certain mechanisms of dissemination of information as conduits for news and propaganda and potential threats to the state. Consequently, the Internet—and with it cyberspace—is a key mechanism for such dissemination. Other states view the freedom of information as the bedrock of cyberspace interaction. This split in views is not going to be easily overcome. TCBMs do, however, offer a route for progress even on such highly political issues. By building an understanding of both perspectives, approaches and possible "red lines", states can identify and navigate towards areas of common ground. This approach works to avoid slipping into a state of attrition involving entrenched political positions, such as on freedom of information. It allows the foundations to be laid to tackle the key questions such as: what are the military rules of engagement for a conflict in cyberspace?

In addition, there are many questions as to where boundaries and red lines are to be found in the cyber environment. Would a state interpret a state-sponsored attack on its largest commercial bank as an armed attack within the meaning of Article 51 of the United Nations Charter?[7] What is included in a state's critical infrastructure and what is a proportional reaction if attacked? Establishing where these lines lie will result in much clearer recognition by policymakers and military actors of the future realities of state-to-state engagement in cyberspace. The United States has been clear that it considers that the existing international legal structure, including the law of armed conflict, is applicable to cyberspace, but has also stated that it sees a need for further work to be carried out in establishing the principles for reaching "a definitive legal conclusion as to whether a particular disruptive activity in cyberspace constitutes an armed attack triggering the right to self-defense".[8]

Regarding the specifics of self-defence, a clearer understanding needs to be developed on what are considered to be the obligations of states to prevent their territory being used for cyber attacks. Obviously, there is once again a clear difference here between cybercrime and cyberwarfare. The position has been put forward that non-belligerents in a conflict are not obliged to prevent the use of their networks as conduits for offensive purposes under the law of neutrality.[9] Such issues require clarification if effective norms of behaviour for all states—not only those with offensive capabilities—are to be developed.

Current initiatives

The London Conference on Cyberspace

The London Conference on Cyberspace, which took place in September 2011, was instrumental in raising the profile of the steps required to build confidence among international partners. It is clear from the discussions at the London Conference that there is still a range of opinion on the exact nature and definition of cybersecurity. The actual question of defining the term was not directly tackled but a clear split emerged between those who view Internet freedom as a fundamental human rights issue and those who have grave concerns regarding national security risks, information security threats and the use of information as a weapon.[10]

This debate underlines the need to divorce the expression of political ideas in cyberspace from the practical steps needed to develop cybersecurity TCBMs and secure cyberspace in the long term. With regards to the discussion on international security at the London Conference:

> All delegates underlined the importance of the principle that governments act proportionately in cyberspace and that states should continue to comply with existing rules of international law and the traditional norms of behaviour that govern interstate relations, the use of force and armed conflict, including the settlement by states of their international disputes by peaceful means in such a manner that international peace, security and justice are not endangered.[11]

It is important to note that the participants did not consider it timely for legally binding measures to be discussed. The true success of events such as the London Conference is in providing a structured non-formal forum in which such common understandings on the next steps for action and discussion can be agreed. It is hoped that the 2012 and 2013 conferences—hosted by Hungary and the Republic of Korea respectively—will play a similar role.

International code of conduct for information security

A proposal made by China, the Russian Federation, Tajikistan and Uzbekistan for an international code of conduct for information security was first circulated in 2011 in a letter to the Secretary-General of the United Nations.[12] While many disagree over its content, the proposal has been an effective tool for spurring debate.

The proposed code, however, does not detail any recommendations on the creation of norms, TCBMs and definitions but instead is confined to broader statements on the nature of information security and the potential use of information as a weapon.

> Each State voluntarily subscribing to the code pledges:
>
> [...]
>
> (b) Not to use information and communications technologies, including networks, to carry out hostile activities or acts of aggression, pose threats to

> international peace and security or proliferate information weapons or related technologies.[13]

Such a position of limited freedom of information has been consistently opposed by various states, notably the United Kingdom and the United States. At the current stage of norms development in cybersecurity, it would seem that the international community is not yet ready to start work on such a document. However, it once again underlines the need for cross-cutting foundational work on terminology and establishing where both disagreement and common ground are to be found before there can be any hope of progress on more elaborate and politically sensitive topics.

Regional organizations

Regional organizations have a long history of working with TCBMs in conventional security areas. Housing such initiatives in a regional organization framework has many positive aspects. First, such an initiative builds on models and lines of communication already familiar to participating states. Therefore, methodologies that have been successful in other areas have the potential to be transferred over to cyberspace. Furthermore, regional organizations may be better able to respond to regional concerns or requirements—especially if the cyber capacities of their member states are at a similar stage of development. As an example, Organization for Security Co-operation in Europe (OSCE) member states, with support of key actors such as the United Kingdom,[14] are investigating the possibility of establishing a working group focused on developing confidence-building measures for cyberspace. The working group would be established by a decision of the Permanent Council of the OSCE. If successful, it may become a model that can be applied by other regional organizations.

A further example is the agreement between the member states of the Shanghai Cooperation Organization on international information security.[15] This agreement, signed by China, Kazakhstan, Kyrgyzstan, the Russian Federation, Tajikistan and Uzbekistan in 2009, takes important steps forward on building common political positions on information security. One of its most significant contributions is the inclusion of a list of definitions of basic terms. This will help future discussions between states to progress, as all parties will have a clearer idea of the conceptual parameters within which others are operating.

UN Groups of Governmental Experts on Information Security

The UN General Assembly occasionally convenes Groups of Governmental Experts (GGEs) to explore areas of particular concern and make recommendations. Membership in GGEs is usually limited to no more than 15 experts, nominated to be geographically representative. GGEs meet in several closed sessions, attempting to reach consensus. If the group is successful in reaching agreement, the resulting report is submitted to the Secretary-General for consideration.

At the suggestion of the Russian Federation, a GGE on the topic of information security was convened in 2004. The group failed to reach agreement.[16] In 2009 a second GGE was convened, with the mandate:

> to continue to study existing and potential threats in the sphere of information security and possible cooperative measures to address them, as well as concepts aimed at strengthening the security of global information and telecommunications systems.[17]

This group reached consensus. Their 2010 report made the following recommendations:

> (i) Further dialogue among States to discuss norms pertaining to State use of ICTs [information and communications technologies], to reduce collective risk and protect critical national and international infrastructure;
>
> (ii) Confidence-building, stability and risk reduction measures to address the implications of State use of ICTs, including exchanges of national views on the use of ICTs in conflict;
>
> (iii) Information exchanges on national legislation and national information and communications technologies security strategies and technologies, policies and best practices;
>
> (iv) Identification of measures to support capacity-building in less developed countries;
>
> (v) Finding possibilities to elaborate common terms and definitions relevant to General Assembly resolution 64/25.[18]

The General Assembly has agreed to convene a new GGE in 2012, with the mandate to take "into account the assessments and recommendations contained in the [2010] report, to continue to study existing and potential threats in the sphere of information security and possible cooperative measures to address them" as well as "relevant international concepts aimed at strengthening the security of global information and telecommunications systems".[19] Based on the structure of the 2010 recommendations, it would seem that the next step would be to develop some specifics on what the implementation of the recommendations might look like and—equally importantly—designate a forum for future discussion of TCBMs in cyberspace.

Forum of Incident Response and Security Teams

As mentioned above, TCBMs on international security and cyberwarfare do not necessarily need to be constructed at the state level. Given the extensive involvement of the private sector in the development of cyberspace, it can also play a major role.

The Forum of Incident Response and Security Teams (FIRST) network is an example of such an undertaking. FIRST is an international confederation of computer emergency response teams

(CERTs), formed in 1990, with the aim of counteracting challenges arising from, for example, differences in language, time zones and international standards. FIRST aims to coordinate CERTs and cooperatively handle computer security incidents and promote incident prevention programmes. Bringing together the educational, government, military and commercial sectors, it provides a mechanism for the coordination of cyber incident response and provides access to best practices and tools, and to trusted communication with member teams.

Such initiatives, while originating from a very specific need, contribute greatly to the internationalization of best practices in cybersecurity. This is of special relevance for states with less capacity in cybersecurity. It is imperative that the international security community looks to mechanisms such as these and ensures that governmental action at the multilateral level is harmonized with the activities of operators and other stakeholders, such as private businesses relying on cyberspace infrastructure.

International Telecommunication Union

With its Global Cybersecurity Agenda, the International Telecommunication Union (ITU) has continued to build a role in cybersecurity and has generally taken a holistic view on the issues of cyberconflict and cybercrime. Hamadoun Touré, the ITU Secretary-General, has outlined five key principles for "cyberpeace":

1. Every government should commit itself to giving its people access to communications.
2. Every government will commit itself to protecting its people in cyberspace.
3. Every country should commit itself not to harbor terrorists/criminals in its own territories.
4. Every country should commit itself not to be the first to launch a cyber attack on other countries.
5. Every country must commit itself to collaborate with each other within an international framework of co-operation to ensure that there is peace in cyberspace.[20]

While these principles are Touré's personal views, they do seem to reflect the general direction of ITU involvement in cyberspace. The ITU should be commended on its continued efforts to set standards, provide capacity-building and build linkages from cybercrime to cyberconflict. Understanding how such procedural and technical work can contribute to the larger, highly political, international cybersecurity debate demands further examination and an understanding of how this best relates to other ongoing initiatives.

Conclusion

The international community is currently at a turning point in cybersecurity diplomacy. States have become aware of the threats and challenges they now face in an environment that is constantly evolving. Given that the initiatives discussed here are still at the early stages of development, the point has not yet been reached when states are politically chained to a particular initiative and thus are not prepared to consider alternatives. This should be taken advantage of. Currently, there is a window of opportunity to make real progress on definitions and operational TCBMs. While no state's concerns should be disregarded, it is imperative to disassociate those measures that are beneficial to all parties at a foundational level from more conceptual questions regarding the balance between information warfare and freedom of expression.

Clarifying military and political doctrine on issues such as the protection of critical infrastructure and national positions on thresholds for a state to take offensive or defensive action in cyberspace provides plenty of substance for working towards near-term progress.

In terms of specific mechanisms for TCBMs, every option should be considered given that the goal is a cyber environment that is more stable, more predictable and less likely to result in miscommunication leading to conflict escalation. Bilateral understandings between advanced Cyber Powers, a multilateral accord or an international private–public agreement all are possible avenues for progress and deserve further investigation into their feasibility.

2012 through 2014 will be crucial years for setting the future direction of the interaction of states on and in cyberspace. TCBMs developed during this period, it is hoped, will work to ensure that the interaction is as peaceful as possible.

Notes

1. For more information on the Stuxnet attacks see R. Langner, "Cracking Stuxnet: A 21st-Century Cyber Weapon", <www.ted.com/talks/ralph_langner_cracking_stuxnet_a_21st_century_cyberweapon.html>.
2. J. Lewis and K. Timlin, "Cybersecurity and Cyberwarfare: Preliminary Assessment of National Doctrine and Organization", UNIDIR, 2011.
3. Ibid., p. 3.
4. See J. Robinson, "The Role of Transparency and Confidence-Building Measures in Advancing Space Security", *European Space Policy Institute Report 28*, 2010, pp. 14–26.
5. General Assembly, *Special Report of the Disarmament Commission to the General Assembly at its Third Special Session Devoted to Disarmament*, UN document A/S-15/3, 28 May 1988, pp. 28–33.
6. The comments made by M. Markoff at the conference International Engagement on Cyber, Georgetown University, 29 March 2011, can be found at <www.acus.org/event/international-engagement-cyber-establishing-international-norms-improved-cyber-security/panel-4-transcript>.
7. Article 51 opens with: "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence 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".

8. US Department of State, *Cyber security keynote address by Dr. Deborah Schneider, US Department of State*, document FSC-PC.DEL/30/10, 9 June 2010, p. iv.
9. See N. Melzer, "Cyberwarfare and International Law", UNIDIR, 2011, § IV.
10. For further information on the Internet and human rights see The White House, "VP's Remarks to the London Cyberspace Conference", speech by US Vice-President Joe Biden, 1 November 2011.
11. Foreign and Commonwealth Office, *London Conference on Cyberspace: Chair's Statement*, 2 November 2011.
12. General Assembly, *Letter dated 12 September 2011 from the Permanent Representatives of China, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General*, UN document A/66/359, 14 September 2011.
13. Ibid, p. 4.
14. "Meanwhile the UK will work actively in the UN and with organisations such as the Organisation for Security and Cooperation in Europe (OSCE) to develop practical confidence-building measures to reduce the risk of escalation and avoid misunderstandings between states arising from unexpected incidents in cyberspace"; Cabinet Office, *The UK Cyber Security Strategy: Protecting and Promoting the UK in a Digital World*, 2011, p. 26.
15. Shanghai Cooperation Organization, *Annex I to the agreement between the governments of the member states of the Shanghai Cooperation Organization on cooperation in the field of international information security*, 16 June 2009, based on an unofficial translation.
16. In addition to the Permanent Five on the Security Council—China, France, the Russian Federation, the United Kingdom and the United States—Belarus, Brazil, Germany, India, Jordan, Malaysia, Mali, Mexico, the Republic of Korea and South Africa were also represented. For further information see T. Maurer, "Cyber Norm Emergence at the United Nations: An Analysis of the Activities at the UN Regarding Cyber-Security", *Explorations in Cyber International Relations Discussion Paper 2011–11*, 2011.
17. General Assembly, *Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN document A/65/201, 30 July 2010, p. 5.
18. Ibid., p. 8. The GGE was composed of the Permanent Five, Belarus, Brazil, Estonia, Germany, India, Israel, Italy, Qatar, the Republic of Korea and South Africa.
19. General Assembly, *Developments in the field of information and telecommunications in the context of international security*, UN document A/C.1/66/L.30, 14 October 2011.
20. ITU, *The Quest for Cyber Peace*, 2011, p. 103.

Achieving mutual comprehension: why cyberpower matters to both developed and developing countries

John B. Sheldon

Cyberspace and the security issues pertaining to it have recently taken up a significant portion of the diplomatic and international security agenda. This is largely due to the perceived and actual threats posed by cyber capabilities to the national security and economic well-being of states. So far much of the agenda and its attendant debates have revolved around the concerns of developed countries, while developing countries have either been silent, ignored or, in some cases, cast as "cyber villains". I contend that while there is indeed a gap in developing world participation in the global dialogue and debate on cyberspace issues, there are in fact several important areas of interdependent and common interest in cyberspace of equal importance to both the developed and developing world.

Characterizations and discussions of cyber threats and so-called "cyberwar" largely revolve around the security concerns of developed countries. While some of their concerns are legitimate, there is a great deal of scaremongering as to the scope and consequence of these threats coupled with an under-appreciation by many policymakers of the difficulties involved in significant cyber attacks. In reality, catastrophic cyber attacks, where power grids crash nationwide, airplanes fall from the sky and financial networks are disrupted, while certainly possible, are unlikely due to the complexity of the targets, the numerous stages in which an attack can fail, and the improbable luck required on the part of the attacker. Just as importantly, developed countries enjoy a measure of redundancy in their military and critical infrastructures that allows for a semblance of continuity of operations in the event of a successful cyber attack against them.

Although some developing countries have been identified as sources of cyber attacks, discussions and debates about cyber threats tend to be silent on the interests of, and possible impact on, the developing world. The threats posed by cyber attacks are just as salient—and in some cases more so—to developing countries as they are to developed ones, yet it is the latter that set the international cybersecurity agenda. However, they both have significant interests at stake in cyberspace.

While the adoption of information and communication technologies (ICTs) by developing countries has undoubtedly provided a perceived military advantage and many benefits and improvements in public services and overall quality of life, it is not an unalloyed good in itself. The same is true in regard to the ubiquity of cyber technologies throughout the developed world. For all the benefits that ICTs have brought, states are becoming equally vulnerable to potentially catastrophic cyber attacks. I identify five key issues that are of mutual and

John B. Sheldon is Professor of Space and Cyberspace Strategic Studies at the School of Advanced Air and Space Studies. The opinions expressed in this article are the author's own and do not necessarily represent the views of the US Air Force or the United Nations.

interdependent concern to all states, and that may provide the basis for dialogue and debate which in turn might hopefully lead to more tangible diplomatic initiatives.

Cyberspace and information threats

Before proceeding to the main discussion, a description of cyberspace and which kinds of cyber threats are of the greatest concern is warranted.

Cyberspace is defined here as:

> [A] global domain within the information environment whose distinctive and unique character is framed by the use of electronics and the electromagnetic spectrum to create, store, modify, exchange, and exploit information via interdependent and interconnected networks using information-communication technologies.[1]

Effects generated from cyberspace—be they strategic (diplomatic, military) or even criminal—are called cyberpower, which is defined as "the ability to use cyberspace to create advantages and influence events in all the operational environments and across the instruments of power".[2] The operational environments comprise land, sea, air and outer space; and the instruments of power are diplomacy, information, the threat and use of military force, as well as economic, social and cultural instruments.

Note that here cyberspace is described as a "global domain" and not a global commons. This distinction is vitally important, as the above description accurately captures the fact that sovereignty can—and is—asserted in cyberspace because the vast majority of the physical infrastructure that underpins it is owned by private corporations registered in sovereign states, with the remaining infrastructure state-owned. This infrastructure includes all of the transoceanic undersea cabling and satellites that route information packets throughout this global domain.

It should also be noted that the definition of cyberspace used here subordinates this global domain to a wider "information environment", also known as the "infosphere".[3] The infosphere is where information flows from all sources—fellow humans, printed materials, radio, television, film and video, as well as cyberspace. In other words, while cyberspace is rapidly becoming one of the more dominant forms of information flow, it is far from the only means available to individuals, private corporations, non-governmental organizations and governments.

Cyberspace has a number of characteristics that explain its saliency to international security policy and debates as well as its growing importance as a strategic domain. These characteristics have both positive and negative implications for individuals, organizations of various types, states and international politics. From a technological perspective the offence–defence distinction is harder to make in cyberspace, but offensive behaviour and actions can

be discerned even though they can be devilishly difficult to attribute in terms of exact source, location and motive.

These characteristics include:

- low barrier for technological entry due to the prevalence and affordability of the required equipment;
- minimal technical skills required, for the most part, to use cyber-technologies;
- growth in the use of cyber-technologies that in turn is leading to near global ubiquity;
- rapid dissemination of data through cyber-networks;
- data can be easily replicated and thus almost impossible to destroy;
- cyberspace is utterly reliant on the electromagnetic spectrum; and
- cyberspace can be stealthy.[4]

The benefits of cyberspace include greater economic productivity and involvement in both the developed and developing world. Indeed, the so-called BRICs (Brazil, the Russian Federation, India, and China—shorthand for rapidly emerging markets) owe their rapid rise largely to the adoption and productive use of cyber technologies.[5] In the developing world mobile phones (or more specifically smartphones, which can access the Internet) are creating new means of personal connectivity and new economic opportunities.[6] Other benefits include the wide dissemination of information for education, better governance and decision-making in general, as well as the automation of various labourious, but necessary, functions that keep the infrastructure of societies functioning on a day-to-day basis.

Of course, there are also problems. While cyber technologies increase connectivity, efficiency and productivity, they also cause tremendous "creative destruction" by eliminating some vocations and creating new ones, which is radically changing the nature of the global economy and as a result creates both winners and losers and the attendant social problems that come with such great change.[7] Cyber technologies are also being used for nefarious purposes, such as a range of criminal activity, recruitment and funding of terrorist organizations, and—given cyberspace's ubiquity, stealth and increasing connectivity to critical systems that maintain day-to-day societal operations—are rapidly becoming a means of surreptitiously stealing from, spying on and even attacking individuals, corporations and governments.[8]

These problems are of the greatest concern to a growing number of states and are now the subject of ongoing discussion. While there is agreement that cyberspace has many great benefits, cyber technologies in the wrong hands pose a serious threat to the national security of individual states—and as a result there are implications for international security. These cyber threats emanate from individuals adept at manipulating cyber technologies, non-state actors, as well as from state entities. Cyber technologies can be used for a variety of malign purposes, ranging from denial-of-service attacks that overwhelm business and government websites with so many requests that they become inoperable—often as an act of political protest or criminal vandalism—to the development and propagation of malware (malicious

software) that infects critical systems and processes for the purposes of espionage, disruption or destruction of cyber-networks. The most well-known case is the Stuxnet virus, which allegedly infected the centrifuge control system at the nuclear facility located at Natanz, Iran.[9] Other threats include logic bombs, which are surreptitiously inserted into target networks in order to disrupt or destroy it at a predetermined time, and social engineering operations, which seek to gain access to computer networks by exploiting the psychological vulnerabilities of users.

These various cyber tools are used to conduct espionage against individuals and corporate, non-state and state entities by accessing proprietary systems that contain sensitive or classified information, and are also used by criminals to steal virtual assets. They are used to disrupt the use of networks and supervisory control and data acquisition systems that provide automated control and feedback of everything from manufacturing processes to the functioning of many elements of modern critical infrastructures, causing physical damage and potentially catastrophic outcomes.[10]

For most developed (and mainly Western) countries the greatest cyber threats are espionage, disruption and crime. Network security and assuredness is generally considered to be the main concern, though the veracity of the information that transits these networks is also considered important. For states such as China, the Democratic People's Republic of Korea, Iran, the Russian Federation, the Syrian Arab Republic and other (largely developing) countries, network security and assuredness are also important. However, equally significant is the concept of information security that emphasizes the veracity of information according to its deemed political and cultural suitability, not just its provenance and incorruptibility. The topic of information security is among the most divisive on the international cyberspace agenda, and it is unlikely to be resolved due to the fundamental and diametrically opposed political and philosophical issues.[11]

With cyberspace defined, the threats described and the most divisive issue in international cyberspace debates identified, I now focus on how developed and developing countries can achieve mutual comprehension of cyberpower issues.

Achieving mutual comprehension

Cyberspace has created interdependent interests and vulnerabilities for developed and developing countries and these should be justification enough for greater dialogue and debate among all concerned. It should be noted that improved mutual comprehension on cyberpower issues, while welcome, does not in itself resolve mutual distrust and hostility that may exist between countries—only committed and skilled statecraft tailored to the particular context at hand can help ameliorate such conditions. However, the issues outlined below are of common concern and interest to all states and, other differences aside, may prove to be

useful "springboard" topics for productive and meaningful dialogue and debate. Hopefully, such conversations may highlight points of comity rather than discord.

There remains, however, the problem stated at the beginning. The diplomatic dialogues and intellectual debates on cybersecurity and cyberconflict are dominated by the major powers and their security experts and defence intellectuals. Conspicuous by their absence in these debates, with a few notable exceptions, are the expressed views and concerns about cyberpower issues from developing countries. Why this should be is a matter of conjecture, although it might be explained by a lack of indigenous cyber expertise, concerns about information as a politically threatening instrument, or scarce human and financial resources being allocated to far more pressing problems. Another plausible reason for the overall absence of developing countries from cyber-themed discussions and debates is that many might be quite content to let the major powers slug it out, though such an approach is certainly not without its risks for all concerned.

Whatever the reasons might be for the general lack of participation by developing countries, there are a number of compelling reasons why developed and developing countries should strive to cultivate greater mutual comprehension of cyberpower issues and, as a result, hopefully create opportunities for more tangible diplomatic outcomes.

The near ubiquity of cyberspace

Cyberspace is rapidly approaching ubiquity. Even in those parts of the developing world where modern cyberspace infrastructure is either non-existent or modest in size and complexity, people are increasingly utilizing ICTs to enhance and improve their social, economic and political interactions. Because ICTs and cyber-networking technologies are cheap and plentiful, their rapid dissemination and adoption in all parts of the world is only going to increase.

As a result, all states have to pay attention to the spread of cyber technologies because their uses are having a profound impact on societal interactions, economic development and activity, political discourse, national security and governance. To be clear, it is not claimed here that cyber technologies are changing the nature of things. Rather, these technologies are changing the very character of day-to-day human interactions and how governments legitimate their claims to power and sovereignty. Cyber technologies that enable the rapid rise and spread of cultural and social trends can also do the same for social and political grievances.[12] For example, online social networks did not cause the uprisings in the Arab world that began in Tunisia in late 2010, but they certainly played significant roles in accelerating awareness and dialogue among protestors on one hand, and in providing critical intelligence about those very same protestors to savvy security services. This phenomenon has put to rest Western notions of cyberspace possessing magical democratizing qualities that cause oppressive regimes to collapse.[13]

Supply and demand

Most cyber technologies are invented, developed and owned by cyber-adept states, and as a result these states are able to have a powerful influence on technical standards, protocols and control of the technological development of cyberspace. This also means that more cyber-adept states are better able to influence the diplomatic and political agenda on cyberspace issues.

This state of affairs, however, does not mean that developing countries should feel that they are automatically excluded. While cyber-adept states may hold proprietary power over the technologies and thus influence over the standards and protocols used, these same states are also eager to sell cyber technologies to the global market. The developing world forms a significant part of that global market and is therefore able to exercise influence on the technologies that are acquired and how they are used—either by developing and enforcing comprehensive laws governing cyberspace within their sovereign territory or by actively seeking to deny their citizens access online to what might be deemed undesirable information. The latter case goes to the heart of the information security controversy mentioned earlier. Most authoritarian regimes are not cyber-adept, lacking the expertise and industrial capacity capable of developing and producing their own cyber technologies. Such regimes, however, are able to purchase software and technologies capable of monitoring who is accessing "undesirable" information available online and even blocking that access. Ironically, many of these censoring technologies are produced and marketed by companies from developed, liberal–democratic states.[14]

Furthermore, while developed countries possess, for now, a near monopoly on the design, development and ownership of cyber technologies, their manufacture and assembly largely takes place in developing countries because of cheaper labour and, in many cases, lax regulations governing issues such as labour or environmental protection. Furthermore, many of the raw materials required for the manufacture of cyber technologies, such as rare earth elements, are predominantly found in the developing world.

Both developed and developing countries have every reason to foster greater dialogue on cyberpower issues given their interdependent market demand and supply chain relationships.

Asymmetry and vulnerability

Developing countries are becoming just as vulnerable to cyber disruption as developed countries, which are deemed today to be more cyber-dependent. This is due to the near ubiquity of cyber technologies, coupled with the fact that these technologies are increasingly used to control critical processes and systems vital to the day-to-day functioning of societies, including those in the developing world.

For example, the growth of so-called "smart" cities in the developing world, coupled with the urbanization of approximately 70% of the world's population by 2050,[15] means that there will be a corresponding increase in vulnerability to cyber disruption in developing countries. Initially, such disruptions may not have catastrophic consequences, but as increasing numbers of critical processes and systems are automated by cyber technologies, the greater the prospect that such disruptions will become more challenging over time.[16] It is imperative, therefore, that developing countries become increasingly involved in international discussions and debates on cybersecurity and cyberconflict in order to mitigate the worst possible effects of these cyber disruptions.

There is also the challenge of many states viewing cyberwarfare as a possible means of gaining asymmetric advantage against more militarily powerful rivals. A number of developing countries may be both tempted to develop cyberwarfare capabilities, and fearful of falling victim to cyberwarfare conducted against them. Despite several isolated incidences of cyber attack (Estonia in 2007 and Georgia in 2008), little is known about the extent and limits of cyberwarfare dynamics between belligerents. Furthermore, while there is much theoretical speculation about the possible consequences of cyberwarfare, the reality is unknown and so the first to engage in such warfare face risks that may backfire on all involved.[17]

Another factor to consider is that cyberwarfare may not confer the asymmetric advantage some might hope for. Cyber attacks against critical infrastructure and command and control networks may achieve some measure of effect, but are very unlikely to result in the immediate capitulation of the adversary. Rather than rolling over in the wake of such attacks, an adversary may well seek to respond by escalating to kinetic means of attack. Warfare by solely cyber means is very unlikely to occur, as the nature of war tends towards extremes. It is, perhaps, more accurate to speak of the use of cyber technologies in war, and those seeking to use it as their primary means of war may find that the outcome is unlikely to be in their favour.[18]

Most developed countries have largely ceded control of complex processes and systems essential to the day-to-day functioning of their societies to cyber technologies and depend on the interconnectedness provided by cyberpower for the functioning of both society and the economy and, as a result, for national security.

Where cyberspace is approaching ubiquity in the developing world, it is not implausible to argue that this has already been achieved in the developed world. Human existence will not end if catastrophic failures in cyberspace were to suddenly occur, but life would be difficult for the affected societies. Everything from financial transactions and telecommunications to critical infrastructure and modern military power are now all critically enabled by cyberspace. This cyberpower bequeaths tremendous advantages and power to those who depend upon it, but it also creates critical vulnerabilities that might prove catastrophic if exploited.

For developed cyber-dependent countries the greatest threats are similarly cyber-dependent countries who are more likely to possess the technical skills and capabilities than developing

countries with more modest resources. Nevertheless, developed cyber-dependent countries might feasibly be surprised by a cyber attack from a developing country that believes it might possess an asymmetric advantage capable of crippling a more developed adversary with a few well-aimed and timely cyber attacks. In reality, such a scenario is unlikely but not impossible. Therefore, developed cyber-dependent countries should seek to engage with the developing world on cybersecurity issues in order to better protect themselves against nasty surprises and seek ways to lessen any cyber threats.

Encouraging global engagement in cyberpower dialogues and debates is vital, therefore, in order for all states to better understand the risks of cyber technologies in war and achieve a greater awareness of adversary "red lines" and appropriate responses to a cyber attack.

Non-state actors

The rise of global organized crime and terrorist organizations equipped with sophisticated technology, such as cyberpower capabilities, should be of grave concern to all states. These entities not only carry out criminal acts that impact lives, the rule of law and even national security, their acts are also capable of undermining the viability and legitimacy of sovereign states. Criminal organizations are exceptionally adept and sophisticated in their use of cyberpower, albeit for nefarious purposes, while terrorist organizations have so far confined their use of cyberpower to recruitment and fund-raising. Yet with the increasing commodification of highly sophisticated and widely effective cyber capabilities (in many cases developed and sold by criminal organizations), there is a very real prospect that terrorists might widen their exploitation of cyberpower. Such actions might include repeated attacks against critical infrastructure that overwhelm a government to such an extent that it loses its legitimacy to rule in the eyes of its citizens, or sophisticated social engineering and propaganda cyber campaigns that suborn influential figures or blackmail senior officials. There is also the risk of states using such non-state actors as proxies to carry out deniable cyber attacks against others.

Both developed and developing countries have very strong incentives to engage in meaningful dialogue on the issue of cyber-enabled non-state actors. In the case of cybercrime some progress has indeed been made with the Convention on Cybercrime, which was drafted by the Council of Europe and came into force in 2004. Discord exists, however, because some states have used the legitimate threat of cybercrime and cyberterrorism as a cover to suppress what many regard as legitimate political protest and dissent.

You may not be interested in cyberpower …

… but cyberpower is interested in you. Finally, even if the preceding arguments fail to persuade, policymakers are left to ponder the fact that even if they are uninterested in the importance of achieving mutual comprehension of cyberpower in global politics, cyberpower

is very much interested in them. States that fail to engage in cyberpower dialogues and debates may truly find themselves the helpless victims of cyberpower wielded by those who grasp its strategic potential even if they do not fully understand the implications of its use.

Conclusion

Cyberpower capabilities are rapidly spreading around the world with often disturbing societal, economic and national security implications. The absence of meaningful developing world involvement in global diplomatic dialogue and intellectual debate on cyberpower issues is an issue of concern for all, and every effort must be made to bring developing world perspectives and concerns to the fore in order to strive for a fundamental level of mutual comprehension about cyberpower, what it portends and how it can be used for the common good. The five points of interdependent interests and vulnerabilities are proposed as a good place to start in closing this dangerous gap.

Notes

1. D. Kuehl, "From Cyberspace to Cyberpower: Defining the Problem", in F. Kramer, S. Starr and L. Wentz (eds), *Cyberpower and National Security*, 2009, p. 28.
2. Ibid., p. 38.
3. See the chapter "Information Power: Strategy, Geopolitics and the Fifth Dimension", in D. Lonsdale, *The Nature of War in the Information Age: Clausewitzian Future*, 2004, pp. 179–200. On information and its central importance to human affairs, see J. Gleick, *The Information: A History, a Theory, a Flood*, 2011.
4. J. Sheldon, "Deciphering Cyberpower: Strategic Purpose in Peace and War", *Strategic Studies Quarterly*, vol. 5, no. 2, 2011, pp. 95–112.
5. J. O'Neill, *The Growth Map: Economic Opportunity in the BRICs and Beyond*, 2011.
6. For further information see International Telecommunication Union, *Confronting the Crisis: Its Impact on the ICT Industry*, 2009, p. 75.
7. See E. Brynjolfsson and A. McAfee, *Race Against the Machine*, 2011.
8. For comprehensive albeit US–centric overviews of cyber threats, see: R. Clarke and R. Knake, *Cyber War: The Next Threat to National Security and What to Do About It*, 2010; and J. Carr, *Inside Cyber Warfare*, 2010.
9. J. Sheldon, "Stuxnet and Cyberpower in War", *World Politics Review*, 2011.
10. On the various methods of cyberspace exploitation, see J. Carr, *Inside Cyber Warfare*, 2010.
11. General Assembly, *Letter dated 12 September 2011 from the Permanent Representatives of China, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General*, UN document A/66/359, 14 September 2011, pp. 3–5.
12. See C. Shirky, *Here Comes Everybody: The Power of Organizing Without Organizations*, 2008.
13. Evgeny Morozov does a useful and timely demolition job of what he calls "cyber-utopianism", in E. Morozov, *The Net Delusion: The Dark Side of Internet Freedom*, 2011.
14. See I. Poetranto, "Behind Blue Coat: Investigations of Commercial Filtering in Syria and Burma", posted on 9 November 2011, The Citizen Lab.
15. General Assembly, *Implementation of the outcome of the United Nations Conference on Human Settlements (Habitat II) and strengthening of the United Nations Human Settlements Programme (UN–Habitat)*, UN document A/65/316, 20 August 2010, p. 4.

16. For further information see S. Smith, "Code is Culture", *Discontinuities*, posted on 15 June 2011, Current Intelligence.
17. C. Demchak, *Wars of Disruption and Resilience: Cybered Conflict, Power, and National Security*, 2011.
18. J. Sheldon, "Deciphering Cyberpower: Strategic Purpose in Peace and War", *Strategic Studies Quarterly*, vol. 5, no. 2, 2011, pp. 95–112.

Confidence-building and international agreement in cybersecurity

James Andrew Lewis

The global digital network has become the backbone of the world economy and a significant new venue for attack, but there has been little progress in negotiation or dialogue in the broader context of international security. The secure use of cyberspace has become a vital national interest of all states. As a result, states believe that malicious activity in cyberspace creates real risk to their security and they fear that, through misperception or miscalculation, such malicious actions could trigger damaging military conflict. This creates strong international pressure for multilateral agreement, but the discussion is at a very early stage.

Most advanced militaries have cyber attack capabilities and many others are acquiring them. We can regard cyber attack capabilities as just another mode of attack, which like a missile or an aircraft can strike the enemy from a great distance. And like aircraft or long-range missiles, cyber attack can serve both tactical and strategic purposes. Cyber attack will not be decisive; cyber attack by itself will not win a conflict, particularly against a large and powerful opponent. But it does provide military advantage and therefore will be used. How and when it will be used can still be shaped by international negotiation. It remains an open question as to how this new aspect of warfare will fit into the existing framework governing interstate conflict, and where modification or new agreement is required to better manage conflict and risk.

A June 2011 UNIDIR report prepared using open-source information reviewed policies and organizations in 133 states and found 33 states that include cyberwarfare in their military planning and organization.[1] These range from states with very advanced statements of doctrine and military organizations employing hundreds or thousands of individuals to more basic arrangements that incorporate cyber attack and cyberwarfare into existing capabilities for electronic warfare.

Common elements of military doctrine include the use of cyber capabilities for reconnaissance, information operations, the disruption of critical networks and services, for cyber attacks, and as a complement to electronic warfare and information operations. Some states include specific plans for information and political operations. Cyber attack blends the techniques and tactics of electronic warfare and signals intelligence. Cyber attack will seek to disrupt opponent command and control, increasing the Clausewitzian "fog of war" by creating uncertainty and by damaging data and communications. A skilled opponent could damage or destroy critical infrastructure—currently only a few major "Cyber Powers" have the capability to use software

---

James Andrew Lewis is a senior fellow and Program Director at the Center for Strategic and International Studies (CSIS). Before joining CSIS, he worked at the US Departments of State and Commerce as a Foreign Service Officer and as a member of the Senior Executive Service. He was the Rapporteur for the 2010 United Nations Group of Governmental Experts on information Security. His current research examines strategic competition and technological innovation. Lewis received his PhD from the University of Chicago.

commands sent over the internet to cause physical destruction but another 30 states are developing military capabilities and non-state actors will gain this ability as techniques and tools are commoditized.

The military use of cyber attack is not the most pressing problem for international security, but it is linked to other malicious behaviours and, in some ways, it offers the easiest approach to agreement, given the many applicable precedents in international security. The more difficult problems revolve around the use of cyber techniques for intelligence purposes and engagement with non-state actors. Both issues, however, fall within the ambit of state responsibilities (although the linkages are not yet well-defined), meaning that it is possible to develop measures and norms that limit risk. The effect of norms can be reinforced by confidence-building measures (CBMs), actions taken between states to prevent or reduce ambiguity, doubt and suspicion and improve international cooperation. Norms and CBMs to increase stability in the military use of cyberspace could reduce the concern shared by many states over the potential for cyberwarfare. Common understandings among states about cyberconflict increase the likelihood of deterring malicious action and they also allow for tacit communication in the event of a conflict with an opponent. Developing such understandings would make cyberconflict easier to prevent or manage.

There is a shared perception among governments that the threat of cyberconflict is escalating out of control—this explains the explosion of national cyber strategies as more than 70 states develop plans and organizations to reduce risk. There is a new willingness to approach the problem of international cybersecurity as an issue that states can manage using established tools of negotiation and agreement. But translating a shared fear into concrete action has proven difficult. Cyberwar has only recently been considered an issue for international discussion despite more than a decade of breathless media accounts of Pearl Harbors and Armageddons. Before 2000, only a few states had just begun to develop attack capabilities, the potential damage from such attack was limited, and these military programmes were highly classified. This was in contrast to the ongoing and energetic international discussions of internet governance, reflecting both the lack of expertise in the internet community and an inability to perceive potential risks to national interests.

Discussion of international agreement to limit cyberconflict dates from the 1990s, but this discussion got off to a bad start by focusing on a treaty as the means to promote security and stability. Scholars proposed complex legal instruments whose distant ancestor appeared to be the Kellogg–Briand pact of the 1920s, in which states renounced war as an instrument of policy. Also, in the 1990s, the Russian government introduced a draft treaty in the United Nations, in what became a recurring annual exercise that never achieved consensus. While the idea of a treaty attracted support in the General Assembly, it made no progress because of strong opposition from a few western states. The drafts of the treaties were, in any case, unimplementable. How would any state address serious issues in treaty compliance and verification for cyber capabilities? Binding commitments to avoid attack or hostile actions may

be unworkable, if only because potential opponents are unlikely to observe them. Important definitional issues were unresolved, probably because they are unresolvable. A commitment to ban "information weapons" is not very useful if we cannot say what an information weapon is, and efforts to define "cyber weapons" quickly run afoul of the overwhelmingly commercial use and availability of information technologies.

If a cyber treaty makes no sense, neither does a simple extension of the laws of armed conflict into cyberspace. There are areas of ambiguity, including the scale and nature of damage from cyber attack that could qualify as the use of force (an essential prerequisite for action in international law). Some potential uses of cyber attack create uncertainty in meeting the obligations of international humanitarian law for distinction, proportionality and discrimination requirements in identifying legitimate targets. There are yet few precedents for resolving these ambiguities and the result is an increased chance for misperception and miscalculation of cyber actions or the intent behind them that states fear could escalate into more damaging conflict.

These problems continue to hamper international discussion of cybersecurity. In the last few years, however, the situation has begun to improve. While the Internet community and its affiliated organizations remain inadequate as a venue for discussion of the international security aspect of cyberconflict, military and diplomatic agencies in a range of states have identified cybersecurity as central problem. This change reflects the realization in many states that the high-speed global network that forms the basis of cyberspace has become crucial to their economic well-being and national security, and a source of risk to their national security. Haltingly, the international community is moving towards discussion, if not agreement, on the scope, nature and constraints of cyberwarfare.

Alternatives to a formal cyber treaty began to appear as early as 2008. Rejecting formal treaties, these alternatives drew upon the experience of global efforts to control proliferation to develop a generalized model applicable to cybersecurity. Instead of a binding legal commitment, they proposed that states develop norms for responsible state behaviour in cyberspace. Non-proliferation provides many examples of non-binding norms that exercise a powerful influence on state behaviour.

Norms shape behaviour and limit the scope of conflict. Norms create expectations and understandings among states on international behaviour, a framework for relations that provides a degree of predictability in interactions in security, trade or politics. In this context, cybersecurity becomes the ability of states to protect their national sovereignty and advance their national interests. Cybersecurity creates new challenges for international security, as states are bound more closely together and as the perception of "transnational" risk increases, but it is largely a still undefined element in this web of relationships among states.

The idea of a norms-based approach has growing international support and, as in the non-proliferation arena, widespread adoption of norms could pave the way for more formal agreements in the future. In July 2010 a Group of Governmental Experts (GGE) convened by the United Nations Secretary-General was able to produce an agreed report on "Developments in the Field of Information and Telecommunications in the Context of International Security". This was unprecedented; in addition to the inability of a treaty to win consensus, a previous GGE endeavour in 2004 had failed. But the 2010 report itself is only 1,200 words long. In contrast, the first GGE had reportedly produced lengthy and detailed drafts that failed to win consensus. The brevity of the 2010 report was one element of its success (and this is a useful guidepost for future GGEs on cybersecurity), but brevity is also an indicator of the larger problems that hamper building international consensus.

The successful GGE conclusion in 2010 reflected a shared perception among the government experts that the risk of cyberconflict had become a serious threat to international peace and stability and that the absence of international agreement increased the risk of a destabilizing cyber incident that could spiral into a larger and more damaging conflict. The states represented on the GGE were united by a deep concern over the possibility of unconstrained cyberwarfare and how this might escalate out of control into physical violence. They agreed that discussions of norms and rules for the use of force in cyberspace, along with other CBMs, would improve international security and the stability of both cyberspace and the international system.

Winning even limited GGE agreement was difficult. It should be noted however that public accounts from both academic and media sources have largely glossed over significant differences expressed within the 2010 GGE. While the experts agreed on the increasing cyber threat, there was, however, little else where there was common understanding. Some states believe that existing international norms and laws are inadequate for cyberconflict. Other states argue that the existing laws of armed conflict are sufficient for cybersecurity, and are deeply apprehensive of doing anything that would appear to constrain freedom of speech. A central issue, as is often the case in multilateral discussion, is the extent to which states might concede a degree of sovereignty in exchange for greater security.

These differences were not frivolous, but rather reflect deep divisions on how to approach international agreement and very different views on the use of force, the norms that apply to state behaviour, and the sources of risk in cyberspace. In light of these differences, the members of the 2010 GGE were able to reach agreement on five general recommendations for additional action:

> (i) Further dialogue among States to discuss norms pertaining to State use of ICTs [information and communication technologies], to reduce collective risk and protect critical national and international infrastructure;

(ii) Confidence-building, stability and risk reduction measures to address the implications of State use of ICTs, including exchanges of national views on the use of ICTs in conflict;

(iii) Information exchanges on national legislation and national information and communications technologies security strategies and technologies, policies and best practices;

(iv) Identification of measures to support capacity-building in less developed countries;

(v) Finding possibilities to elaborate common terms and definitions relevant to General Assembly resolution 64/25.[2]

These are valuable first steps. Buoyed by the adoption of a consensus report by the 2010 GGE, a few months later the Russian Federation proposed to the First Committee of the General Assembly that a new GGE be established to continue this work. A new group of experts will convene in August 2012. But the discussion of CBMs also faces significant difficulties. Translating the experience of earlier measures applied in other contexts requires effort, if only because the technologies used in cyberconflict are so widespread. The high degree of secrecy that surrounds state cyber activities—a legacy of their signals intelligence heritage—slows any exchange of information. Misunderstandings over the nature of cyberconflict hamper discussion—the frequent resort to nuclear analogies, which are usually inappropriate for cyberwarfare, are examples of this. Much of the open literature descriptions of cyberconflict are imprecise. The combination of a high degree of secrecy and weak research methodology complicate policymaking.

While there is general agreement on a norms-based approach and that cyber norms and CBMs are essential for international security, there is however very little work on specific proposals that would link cybersecurity to the larger international security "system". We still need to define not only an achievable end state for international cooperation in cybersecurity, but also a path to get there. If the objective is to shape state behaviour through a global framework for cybersecurity, there are many intermediate steps yet to be defined. International discussion will need to begin with measures to build confidence and trust.

For some states, the term cybersecurity itself is inadequate. They believe that the issue is "information security". They argue that information is a weapon and that the laws of armed conflict are inadequate for dealing with this new threat to international peace. They have put forward, under the umbrella of the Shanghai Cooperation Organization, a draft Code of Conduct for Information Security intended to shape discussion at the next GGE by blending objectives such as increased law enforcement cooperation with their own concerns about access to information. For the authors of the Code, stability and security are best achieved by giving states sovereign control over the "information space" and by renouncing the threat or use of force in cyberspace.

The fundamental issue for the next GGE is to further elaborate the 2010 recommendations into concrete measures where international agreement could reduce risk from conflict in cyberspace. Accomplishing this requires consideration of both substance and politics, determining both how to achieve restraint and where cooperation is possible now. Common understandings on a range of issues will be essential—these include on how the existing laws of war apply to cyberconflict, on the nature of escalation in cyberconflict and on the responsibilities of states before and during cyberconflict. Shared understandings among states on these topics would help to create an international framework to constrain cyberconflict and to define the potential consequences for differing levels of hostile action. For each of these issues, however, there are ambiguities and, unsurprisingly, there is a wide disparity of views among key states on the nature of the problem.

Challenges

Any future agreement will need to find ways to deal with broad areas of disagreement. There is no agreement on the thresholds for cyberconflict, particularly the key threshold of what constitutes the use of force in cyberspace and justifies the use of force in response. Perhaps most importantly, there are no shared views on the responsibilities of states in cyberspace. This is an unstable environment.

In part, this reflects differing assessments of the sources of risk. Some states see information as a weapon and as much an element of cyberwarfare as "hacking." When a state says that information is a weapon that could be used against them, they are serious—free access to information is seen as a threat to the regime's stability and survival. That this threat is not intentionally (or consistently) directed at them does not lessen it.

The treatment of information is directly linked to the issue of how states will expand sovereign control in cyberspace. The existing governance model, which depends on an almost tribal assembly of stakeholders in various frail institutions, is inadequate for the security and stability needed for a key global infrastructure. Many governments, finding the current situation intolerable, are exploring where it is appropriate for them to increase their role, to reduce risks to their economies, public safety and national security created by a weakly governed internet. The turbulence over internet governance, as states extend sovereign control into cyberspace to protect their national interests, will complicate reaching agreement on norms for international cybersecurity.

There is a wide disparity of views on how to address the problems of cybersecurity. What kind of agreement (implicit or explicit) is needed, what form these agreements should take, their scope and even their venue remain largely undecided. There is agreement that cyber activities are a legitimate military activity, but no agreement on the rules that should apply to it. There is an ambiguous relationship between cyberwar and espionage. This ambiguity increases the

risk of miscalculation or escalation of cyberconflict as there is only a fine line between breaking into a computer to spy and breaking in to attack.

There are key areas of ambiguity in the applicability of existing laws of armed conflict to cyberconflict, including the treatment of third party sovereignty and the amount and nature of damage from cyber attack that could be interpreted as the use of force. Some operational issues, such as the degree of prior assessment needed to meet the requirements of international law for distinction, proportionality and discrimination requirements in identifying legitimate targets, are also unclear. There are as yet few precedents for resolving these ambiguities. While a new GGE might usefully review these ambiguities, it would be ill-advised to seek to resolve them given the limited chance of reaching agreement at this time.

Obstacles to reaching a multilateral agreement

The immense utility of cyber action will shape any international agreement on cybersecurity. States will not give up this new tool for state power. Cyber attack is cheap and offers strategic advantage. First, the importance of information superiority in warfare and the ability to gain real military advantage from the use of information assets makes digital infrastructures too valuable a target to be declared off limits or for cyber attacks to be renounced. The necessary technologies are either commercial or easily derived from widely available commercial products—a laptop computer, an internet connection and a few computer programs. We cannot control the "precursors" for assembling these "weapons". They are cheap, small, portable, easy to conceal and, for sophisticated programmers in or out of government, easy to construct. Special purpose tools for cyber attack are widely available on thriving cybercrime black markets. It is unlikely that any state will renounce the use of cyber attacks.

Nor would a treaty that excludes certain targets from cyber attack make sense. Existing laws of war already define safeguards and limitations on (but do not ban) attacks on civilian targets. We cannot expect more for cyberspace. An alternate approach could be based on non-proliferation, where states developed multilateral norms that define responsible behaviour. The simplest norm would extend existing law and practice to say that a state is responsible for the behaviour of those on its territory—this would constrain the use of proxies and "patriotic" hackers.

Second, action in cyberspace has been an immense boon to espionage. The close linkage to espionage makes states reluctant to discuss or even admit they possess cyber capabilities, and this linkage also makes it unlikely that they will agree to "ban" first use. A "no first use" commitment could require states to renounce cyber espionage—something they are unlikely to do. Since the techniques of attack and espionage are similar, asking for a commitment not to develop or use cyber tools for penetration of opponent networks is really asking for a commitment not to spy. A "no first use" commitment could even be destabilizing if a victim were to misinterpret an instance of cyber espionage as an attack.

The perceived difficulty of attribution of an attack may encourage some states to believe that they can successfully engage in covert cyber action while evading responsibility. A covert attack where the identity of the attacker is unknown has much less political risk. In addition, mercenaries (usually cybercriminals recruited by a state) can launch sophisticated attacks, providing an additional degree of deniability. The difficulty of attribution is often overstated, as it is increasingly possible in many cyber incidents to determine who is responsible using forensic techniques or active intelligence measures, but the perceived attribution problem increases the temptation to use cyber attack.

These problems mean that approaches that seek to limit cyber attack through multilateral agreement on technological constraints face intrinsic and potentially insurmountable difficulties. Cyber attack is a behaviour rather than a technology. Cyberconflict is shaped by covertness, ease of acquisition and uncertainty, and a legally binding convention that depends upon renouncing use, restricting technology, or upon verification of compliance is an unworkable approach for reducing the risk to international security from cyber attacks. An effort to secure an overarching cybersecurity agreement or treaty that attempted to address the full range of cybersecurity issues would be impractical.

An incremental approach

Agreements to reduce the possibility of misinterpretation, escalation or unintended consequences in cyberconflict are a legitimate subject for international agreement and would improve international security. Just as states feel a degree of constraint from norms and agreements on non-proliferation, establishing explicit international norms for behaviour in cyberspace would affect political decisions on the potential risks and costs of cyber attack. The effect of globalization—the deep economic interconnection among states—has if anything increased the need for cooperation among states.

The creation of norms for responsible state behaviour in cyberspace, the expansion of common understandings on the application of international law to cyberconflict, and the development of assurances on the use of cyber attacks would increase stability and reduce the risks of miscalculation or escalation. The single most important norm for multilateral agreement might be a norm that establishes state responsibility for the actions of its private citizens—such a norm could make it more difficult for states to tacitly encourage proxies by ignoring them or denying involvement with their actions.

However, even simple norms face serious opposition. Conflicting political agendas, covert military actions, espionage and competition for global influence form the context for international discussion of cybersecurity. While there is little or no support for the idea of a treaty, and while international efforts now focus on a norms-based approach, the level of distrust among powerful states is too high for easy agreement on norms.

Disparate values and deep distrust shape the environment for negotiation. Fundamental differences over values, despite formal acceptance of universal human rights, means that the initial set of norms likely to be acceptable to many states is limited. Ultimately, increased stability and security in cyberspace will require common understandings among states on their national responsibilities, on how the laws of war apply, where restraint in the use of the new military capability is possible, and where red lines or thresholds for escalation might exist. But there is too much distrust among competitors to move immediately towards global norms for cybersecurity.

This suggests that international efforts should first focus on CBMs as a foundational element in creating stability and security in cyberspace. CBMs, which require agreement on process rather than on values, could be more attainable in the early phase of creating an international framework for cybersecurity. Incremental steps that focus on reaching multilateral agreement on confidence-building processes for transparency and communication —such as increased transparency in doctrine—may be the most productive approach for reaching agreement in the near term.

Judging from recent and valuable discussion at the multinational conference "Challenges in Cybersecurity" (held 13–14 December 2011 in Berlin, and sponsored by the German Federal Foreign Office, Freie Universität Berlin, the Institute for Peace Research and Security Policy at the University of Hamburg and UNIDIR), there is agreement on the benefits of CBMs, although the portfolio of suggested measures is relatively weak. The leading candidates include greater transparency in doctrine, better mechanisms for crisis management, improved law enforcement cooperation and shared understanding on the application of the laws of armed conflict to cyber attacks. Further work to expand and refine confidence-building measures in cybersecurity will be essential for long-term progress.

Notes

1. *Cybersecurity and Cyberwarfare: Preliminary Assessment of National Doctrine and Organization*, UNIDIR, 2011.
2. General Assembly, *Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, UN document A/65/201, 30 July 2010, para. 18.

New publication

Global Nuclear Stability: Building Greater Accountability and Cooperation

Pavel Podvig (UNIDIR, 2011)

The protection of nuclear material and facilities involves a broad range of activities at both the international and state level. International law recognizes that each state has a responsibility for implementing these measures and providing adequate protection for the nuclear material in its possession. At the same time, the international community has established a set of arrangements that help to create and maintain the nuclear security regime.

This book provides an overview of the international agreements, programmes and institutional arrangements that form the core of the international nuclear security regime. It discusses proposals to strengthen accountability arrangements, as well as the challenges of expanding the scope of the regime and creating a framework for global nuclear security efforts. It demonstrates that despite the progress made at the Nuclear Security Summit, further multilateral action will be required to secure nuclear materials and prevent nuclear terrorist attacks.

The publication forms part of the ongoing UNIDIR project International Cooperation Mechanisms on Nuclear Security, which aims to provide policymakers with analyses of challenges and opportunities in the field and help practitioners and policy experts in their efforts to strengthen the international regime to combat the threat of nuclear terrorism.

For more information on this and other publications, please visit our website <www.unidir.org>.

New project

Norms on Explosive Weapons

High-level United Nations officials, including the Secretary-General and the Emergency Relief Coordinator, have repeatedly expressed concern over the effects on civilians of explosive weapons violence in populated areas. Recent responses to the use of explosive weapons in Homs, the Syrian Arab Republic, demonstrate that a growing number of states, international bodies and civil society organizations have come to recognize that the practice constitutes a serious humanitarian problem that needs to be addressed.

The Norms on Explosive Weapons (NEW) project explores laws and policies governing the management and use of explosive weapons at the international and state levels. It analyses how norms protect civilians from the effects of weapons such as artillery shells, air-dropped bombs and improvised explosive devices.

Building on the UNIDIR project Discourse on Explosive Weapons, this project aims to deepen the understanding of the normative aspects of explosive weapons management by states and to clarify when states consider the use of explosive weapons in populated areas acceptable. The NEW project builds on past achievements in the area of humanitarian disarmament and raises further awareness of the great human costs associated with the use of explosive weapons in populated areas and the legal and moral issues involved. It supports efforts aimed at preventing and reducing civilian harm from this type of armed violence, improving the protection of civilians during armed conflict and strengthening applicable legal frameworks. The findings of this research project will be published in mid-2012. To track the project's progress, please visit <http://explosiveweapons.info/>.

For more information, please contact:

Maya Brehm
Tel.: +41 (0)22 917 11 41
Fax: +41 (0)22 917 01 76
E-mail: mbrehm@unog.ch