2012 4th International Conference on Cyber Conflict
C. Czosseck, R. Ottis, K. Ziolkowski (Eds.)
2012 © NATO CCD COE Publications, Tallinn

The Significance of Attribution to Cyberspace Coercion: A Political Perspective

Forrest Hare
Center for Peace and Security Studies
Georgetown University
Washington, D.C., United States
fbh5@georgetown.edu

Abstract: The question of cyber deterrence, or "What and how do you deter malicious actions in cyberspace?" has been hotly debated over the last few years. Stories of massive intellectual property theft and identity theft cases have surfaced in the Western news spurring several seminars and writings on the subject. Unfortunately, the discussion to date has not moved us effectively toward a comprehensive framework for building a coercion strategy. Most importantly, the debate has failed to accurately characterize the coercion challenge. In most cases confronting developed nations, the more pressing issue is not deterring an actor from choosing to conduct hostile intrusions in cyberspace but compelling the actor to stop conducting intrusions that already have been highly successful. Accurately recognizing the existing dynamic changes coercion calculations in several ways, such as the significance of positive attribution – an important component of coercion theory. Although the proposed coercion strategy framework in this paper will necessarily be less than comprehensive, one important outcome will be that the issue of unequivocal attribution may not be as critical as previously suggested.

Keywords: *cyber security, security studies, attribution, coercion, compellence*

"A difficulty with our being an unaggressive nation, one whose announced aim has usually been to contain rather than to roll back, is that we have not settled on any conventional terminology for the more active kind of threat."

THOMAS SCHELLING (1966)

The opinions expressed, or implied, in this article are solely those of the author, and are based on open sources. Under no circumstances may they be correlated to or perceived as the views of Georgetown University, the U.S. government in general, or the Department of Defense in particular.

1. INTRODUCTION

The question of cyber deterrence, or "What and how do you deter malicious actions in cyberspace?" has been hotly debated over the last few years. Stories of massive intellectual property theft and identity theft cases have surfaced in the Western news spurring several seminars and writings on the subject. Unfortunately, the discussion to date has not moved us effectively toward a comprehensive framework for building a coercion strategy. In fact, authors since the mid-1990s have been arguing that traditional deterrence theory is difficult to apply to current cyber threats (Alperovitch 2011; Harknett 1996; Libicki 2009). However, the debate continues because policy-makers remain unable to find efficacious answers to persistent, and immediate, threats in the domain. The aim of this paper is to advance the discussion forward by using a different perspective on coercion. Most previous writings have focused on the difficulties of applying traditional deterrence theory to the domain, such as the challenges to determining attribution. Most importantly, the debate has failed to accurately characterize the coercion challenge. In most cases of cyber conflict confronting developed nations today, the more pressing issue is not deterring an actor from choosing to conduct hostile intrusions in cyberspace but compelling them to stop conducting intrusions that already have been highly successful.

Accurately recognizing the existing dynamic changes coercion calculations in several ways. For example, it may alter the importance of positive attribution—an important component of coercion theory. To provide a different perspective on the significance of attribution, this paper proposes a cyberspace coercion framework that draws on insights from Schelling (1966), and modeling by Byman, Waxman, and Larson (1999) of RAND. The model by Byman et al. identifies a continuum of policy objectives, from deterring an actor from intruding in systems connected through cyberspace, to one of forcing an actor to stop threatening intrusions and remove malware implanted in critical infrastructure. Based on these objectives, the paper will highlight the relative importance of emplacing strong defenses, communicating retaliatory actions, achieving attribution, and executing effective responses to successful intrusions. It will build on the author's previous work (Hare 2010) regarding international cyber security dynamics to explore effective ways to "ratchet up the pain" necessary to compel actors to change their behavior in the domain. The goal of the paper is to revisit the issue of attribution through this framework and re-assess the importance of positive attribution. Though this proposed framework will necessarily be less than comprehensive, one important outcome will be to reveal that the issue of unequivocal attribution may not be as critical as previously suggested by many authors.

Before making the argument for a more appropriate coercion framework, I will establish a definition for the concept of national security in cyberspace that focuses the discussion on issues regarding international security relations. Thereafter I will provide a short review of the attribution problem in cyberspace as it is currently portrayed in the literature. This review will be followed by my argument on the need for a new coercion model for cyberspace. Using the new coercion model, I posit three potential coercive measures that will provide the opportunity to reassess the attribution problem. The paper concludes with points policy-makers should consider regarding attribution, given this new analysis framework.

2. THE CYBERSPACE THREAT TO NATIONAL SECURITY

In this section, I establish a definition for the concept of national security in cyberspace. This definition bounds the cyber security problem to security issues between nation-states. Using Buzan's (1991) concept of securitization and previous work by this author (Hare 2011), I first specify the public good of national security as, "that state in which the public of a nation is not threatened by something, or someone, that poses an existential threat."[1]

There are two primary ways this state of being can be threatened through cyberspace by adversarial nations and other malicious actors. First, a nation can suffer a threat from intrusions through cyberspace by either state or organized non-state actors against government, and select other, information systems to gain knowledge of national security value. Such activity, whether conducted by people intercepting bits and bytes of information or using their own eyes and ears, is generally considered espionage. Targets of such espionage could include the sensitive information systems of defense ministries or contractors that develop major weapons systems. Successful attacks would allow an adversary to counter a wide-array of national defense measures and they could justify governments using extraordinary measures to thwart such attacks, such as calls for increased deterrence options.

Second, a nation can suffer an existential threat from attacks and infiltrations through cyberspace by either state or organized non-state actors to degrade or disrupt critical infrastructure systems, both privately and publicly owned. For example, emplacement of malware and other disruptive software in the control systems used in the energy, transportation, or telecommunications sector could endanger many lives directly or thwart physical actions intended to defend national interests. Successful intrusions or attacks could also have a significant economic impact or cause a loss of life, and therefore again justify extraordinary counter-measures. Adding these two criteria to the definition of national security redefines the definition of the public good of national cyber security as the state of being in which the populace, governing institutions, and critical infrastructure are not threatened by:

- Attacks and intrusions through cyberspace, by either state or organized non-state actors, against government and select other information systems to gain knowledge of a national security value, or
- Attacks and intrusions through cyberspace, by either state or organized non-state actors, against critical infrastructure systems to degrade or disrupt such systems and cause a national security crisis.

This definition provides policy-makers with boundaries within which to develop a potential coercion strategy. An important component of this definition is the list of malicious actors against which a coercion strategy can be directed. Specifically, coercive actions would be taken to influence actors under direct control of the state or those acting with at least the tacit approval of the state. This second category could include paramilitary organizations, and contractors (Lewis 2011). In either case, state institutions of the adversary regime must be able to influence

1 Significant portions of this section are adapted from *The Interdependent Nature of National Cyber Security: Motivating Private Action for a Public Good* (Hare, 2011). For a more in depth discussion of the concepts in this section, please refer to this work.

both the action and inaction of the malicious actor. In other words, if coercion measures implemented by a threatened nation-state are to be successful, they must be directed at the authorities of the adversary nation who, in turn, must be able to exert their sovereign powers within their own territory. Actors such as patriotic hackers, criminals, and terrorists are much more difficult to coerce, as they are less susceptible to national power. However, since they seldom engage in the activities contained in the above definition of national cyber security, they are also a less significant threat to a nation's sovereignty (Lewis 2011).[2] With the problem thus bounded, I will now return to the calculus of coercion as it pertains to this domain.

3. THE CALCULUS OF CYBER COERCION

As mentioned in the introduction, several seminars and writings have been dedicated to a discussion of deterrence in cyberspace. Most have addressed the difficulties of applying deterrence theories in this domain, and many have focused on the challenges of achieving conclusive attribution of the malicious actors. For example, in his book *Cyberdeterrence and Cyberwar*, Libicki (2009) defines cyberdeterrence as an in-kind deterrence against attacks through cyberspace (p. 34). Using this definition, he highlights several issues that make such a strategy difficult to implement. His first question is, "Do we know who did it?" (p. 41). This is, of course, the issue of attribution. In his opinion, the victim must be able to convince third parties that the attribution is correct and, more importantly, the attacker must be convinced that the act will be correctly attributed to them. Libicki provides several examples of the difficulties in achieving conclusive attribution, based on the anonymity provided by the structure of the Internet and the indirect ways packets can be routed to their eventual targets.

Libicki (2009) even questions the idea that the beneficiary of an action would be its most likely instigator (i.e., *cui bono*), in that there are often many parties that could benefit from an attack. For example, several nations would be interested in intellectual property information from more advanced nations, and armed with this fact alone, it would be difficult to determine which nation had been responsible for a theft of intellectual property. Libicki also raises the possibility of false flagged operations being conducted to divert attention away from the malicious actor. Finally, he raises the practical concern that actions taken to demonstrate conclusive attribution to the international community or directly to an accused attacker will do nothing more than instruct the attacker on how to hide their activities more effectively. Clark and Landau (2010), in a paper specifically devoted to the challenge of attribution, state that "attribution is central to deterrence [...] [and] retaliation requires knowing with full certainty who the attackers are" (p. 25). With this imperative, the reader is left to assume that no deterrence strategy, whether intended to combat crime or defend a nation, can be effective without positive attribution.

Boebert (2010) breaks attribution down into technical and human components then discuss the barriers to achieving either forms, such as the proliferation of botnets and onion routing. He likens the problem of human attribution to that faced by any law enforcement agency that tries to solve a crime based on ballistic evidence. How do we prove who was at the keyboard at the time of an attack even if we identify the offending machine (Boebert 2010)? These authors and others have identified additional challenges to a successful deterrence strategy. Examples of other problems with deterrence in cyberspace include the difficulty of communicating a credible

2 A notable exception could be the events in Estonia in 2007 that will be discussed in this paper.

threat, the lack of clear red lines, and the risks of targeting innocent third parties with automatic responses (see Alperovitch 2011; Clark & Landau 2010; Harknett 1996; Libicki 2009; Lukasik 2010; Taipale 2010). With all of these challenges to a developing a robust deterrence policy, it is time to reassess the coercion problem in the cyber domain.

As the above review has shown, cyberspace presents many challenges when applying traditional deterrence theory. Nonetheless, the pressure to discuss deterrence in cyberspace has driven us to keep raising the attribution issue since it is perceived to be so critical to deterrence. To provide another perspective on attribution, I argue that we need to take different look at the problem of coercion: Is the problem really one of deterring an adversary from attacking us in cyberspace or is it a problem of compelling them to *stop* threatening intrusions that have thus far been very successful? There have been at least three instances of successful intrusion events that would support considering a different perspective.

The first such event was a broad intrusion set known as Ghostnet, which was first discovered by researchers in March 2008 and appeared to be continuing more than a year later (Deibert and Rohozinski, 2009). While the motivation and identity of the perpetrators has yet to be conclusively determined, the intrusion activities clearly targeted the communications systems of the office of the Dalai Lama, the Tibetan government-in-exile, and several non-governmental organizations affiliated with the Tibetan community (Deibert and Rohozinski 2009). Researchers with the Information Warfare Monitor identified an elaborate network of control servers and command servers that were being used to deliver and monitor targeted malware and exploit the information contained on over 1,000 computers in more than 100 countries. By the time the targeted communities discovered what was happening to them, there was no chance of deterring the perpetrator from conducting an act of cyber espionage. The problem instead became how to prevent the intrusion from continuing.

In 2011, a defense department official in the United States stated that, over the past few years, crucial files stolen from defense and industry data networks have included plans for missile tracking systems, satellite navigation devices, surveillance drones and top-of-the-line jet fighters (Shanker and Bumiller 2011). Once again, the military official was not aware of a potential threat to the critical data systems until the attack was well under way. The intrusions that the official is referring to may be continuing without their knowledge. At the very least, the intrusions clearly had occurred over an extended period without encountering any appreciable resistance.

Next, consider large-scale, Distributed Denial of Service (DDOS) attacks. In the case of some technologically advanced nations that have limited Internet-bandwidth, a DDOS against systems such as national banks and government communication systems could pose a significant risk to national security. For example, the small European country of Estonia experienced what it considered to be a debilitating series of DDOS attacks in 2007. These attacks occurred over several days, and there was little advance warning to give the nation's cyber defenders an idea of how broad or successful the attacks would be (Landler and Markoff 2007).[3] In this case, Estonia had no opportunity to develop, let alone communicate, a deterrent threat to any potential adversary. One would expect any potential victim to face this same challenge in deterring any

3 I would also argue that the sponsoring attackers probably did not know how broad or successful the attacks were while they were occurring.

potential DDOS attack that was not announced in advance. Given that a DDOS attack can be launched with virtually no warning and that doing so will greatly improve its effectiveness, such attacks can be expected to be used as first-strike weapons by any potential adversary.

Lastly, intrusions on the power grid in the United States have left behind software programs that could be used to disrupt the power production network, according to current and former national-security officials (Gorman 2009). According to a report in the *Wall Street Journal*, the intrusions were not detected immediately by the targeted power companies but by intelligence officials who identified pervasive espionage within the critical infrastructure sector (Gorman 2009). One would expect that several incidents similar to the four mentioned here have occurred but will remain unreported, due to their implications for national security in the targeted countries. A short survey of several government websites indicates that many countries are continually encountering intrusion activity on their government information networks (see, for example, Australian government n.d.). This success carries a message to hostile actors that such malicious activities will continue to be very rewarding, despite any strong rhetoric from victims. Based on the four incidents described here, I argue that we should revise the calculus of coercion. In so doing, we may find that attribution at the technical or legal level envisioned by previous authors may not be as critical as they conclude.

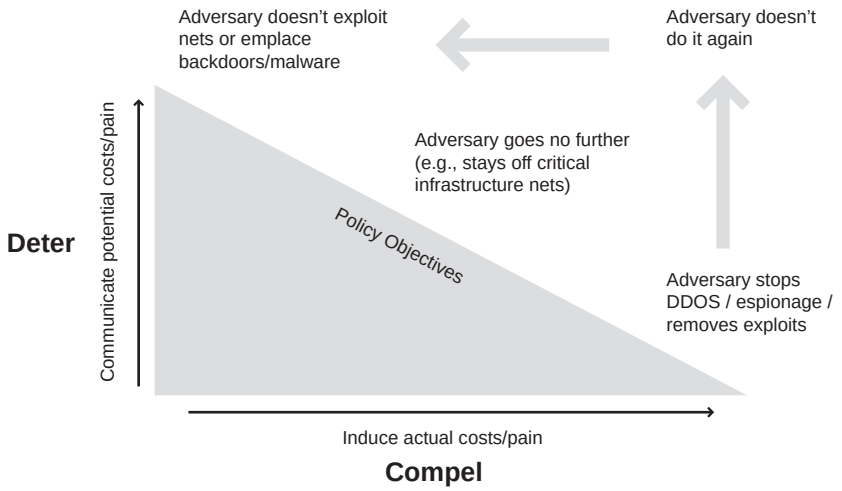
To develop a coercion model applicable to cyberspace based on the above evidence, we should take a fundamentally different approach than that taken by previous authors. If the malicious actors—adversary nation-states or their non-state proxies—have been conducting successful cyber espionage for an extended period of time, have sponsored no-notice DDOS attacks, or have already penetrated critical infrastructure control systems, then this author would argue that the coercion calculus is very different. The situation many developed countries now face is one that has been characterized by Nobel Laureate security strategist Thomas Schelling (1966) as one of *compellence*, not deterrence. According to Schelling, there are important distinctions between deterrence and compellence as components of a coercion strategy. The main differences are in the timing and the initiative. In a compellence situation, the attacker already has accomplished the offending action and the defender must take the initiative to respond, not just sit and wait. In other words, "The threat that compels rather than deters often requires that the punishment be administered *until* the other acts, rather than *if* he acts (Schelling 1966)." The compellence action taken must continue, or be believed to continue, until the offender responds favorably. There is no longer an ambiguous trip-wire that must be triggered before a threatened response is enacted. The line has been crossed, whether or not either party realizes it, and the offender has discovered the benefits have been worth the risk. The impetus is now on the victim to respond with a retaliatory action or assume the increased risk to national security. How much action is necessary will only be clear after the offensive behavior has been reversed or reduced to a level that is no longer considered threatening. However, in a deterrence situation, the defensive picture already has been painted. The adversary need not know the specific features of the painting, as long as no offensive act is committed. In fact, ambiguity may support deterrence. In a compellence situation, the picture must be painted for that specific situation and it must be clear to the offender what must be done, and by when, for the victim's coercive response actions to cease.

This simple change to the dynamic creates its own sets of challenges. First, to retain political

flexibility, it is no longer enough to leave the threshold or response ambiguous. The initiating party must communicate to the adversary that a retaliatory action is being taken in response to a specific action that is deemed hostile and is politically attributed to actors under the adversary's control. Whether or not the responsive action is to be initiated immediately, a deadline for compliance must be clearly articulated so the offender has no question as to when the offending act must cease. Second, not only must the terms be clearly communicated; the communication may have to be done in private so the adversary can avoid the appearance of having to back down. Inaction is easy to justify in a deterrence situation, as a would-be adversary can always claim other reasons for not conducting an action for which a victim threatens retaliation (Schelling, 1966). However, in a compellence situation, it is difficult for the offender to avoid the appearance of bending to the victim's will if the victim is to successfully influence the appropriate change in behavior. Third, in a compellence situation, the victim must develop a retaliatory action that will be effective and executable but that also can be stopped or reversed; otherwise there is no incentive for the offender to cease the offending behavior. Finally, on a positive note, if an effective compellence threat can be emplaced before the damage is too great, it may help overcome challenges faced in communicating the deterrence threat in the first place (Oh, you didn't know that was bad? Well, now you know, and if you stop, so will I). Given these challenges and opportunities, it is extremely important for the policy-maker to understand when they are presented with a compellence challenge.

Figure 1, which is adapted from a model produced by RAND researchers studying the use of air power as a coercive instrument (see Byman, Waxman and Larson 1999), depicts a framework for analyzing the coercion problem in cyberspace.

FIGURE 1. THE COERCION DYNAMIC AND POLICY



This figure shows the relationship between deterrence and compellence for achieving policy objectives to counter hostile cyber intrusions. The X and Y axes depict the relative weights of compellence and deterrence measures that must be taken to achieve exemplary policy objectives

presented along the slope of the triangle. For example, if the policy objective is to dissuade an adversary from executing threatening actions in cyberspace, then it may be sufficient to communicate to them the potential costs of taking such actions. As argued previously, several authors have assessed the many barriers to successfully communicating to the potential attacker that the costs of an offensive cyber action will outweigh the benefits. At the other end of the slope is a pure compellence situation in which the adversary has successfully penetrated and gained significant control over critical networks in a target country, or is currently conducting (or sponsoring) an effective DDOS attack against a victim's critical information systems. Up to this point, assuming the attack has been attempted, mere communication of the costs clearly has been unsuccessful. The only policy choice now available to the victim is to take the initiative to induce actual costs on the attacker in such a way that they respond to the counter measures and alter their offensive behavior.[4] If a victim finds itself in a situation where compellence measures have successfully removed an adversary from sensitive networks, it will most likely need to "keep the pressure on." A return to the *status quo* will be no more likely to deter the resumption of hostile actions by the adversary than it was before. Therefore, a policy mix must continue to induce some actual costs while also threatening the same or even stronger retaliatory measures, should the adversary return to exploit critical networks.

There are a few additional considerations I would like to address before proceeding with the analysis. First, the opaque nature of actions in cyberspace makes it difficult for the defender to know how far the attacker has penetrated and, therefore, exactly where they are on the policy slope.[5] Espionage will exist at some level and in all directions as long as the international system exists. If the victim finds itself in a situation where it sees that the attacker has penetrated to a certain point not viewed as an immediate threat to national security, then the appropriate coercion strategy may be a combination of deterrence and compellence measures. Second, when confronted with a compellence situation in cyberspace, the greatest policy challenge is to identify the appropriate costs or pain to be inflicted on the attacker to make them change their behavior in the desired manner (e.g., to get them off the critical networks). If the policy is restricted to taking retaliatory actions in cyberspace, then the victim's options may be limited. The counter to this point is that there is value in showing connectedness in response. A response that is closely connected in type and degree to the offensive action is easier to communicate to the offending party and to justify to an international audience (Schelling 1966). For example, launching cruise missiles in response to an act of cyber espionage may result in a proportional dollar loss to the offender, but it most likely will not be viewed by many as appropriate or sufficiently linked to the hostile cyber act that provoked the retaliatory measure. Finally, improper or poorly articulated goals can lead to a misapplication of pressure through coercive actions that will neither achieve desired results nor be measurable in any meaningful way. For example, it simply is not possible to stop all malicious actions in cyberspace, at least not with existing technology. However, it may be possible to influence the malicious behavior of nation-state actors to a measurable degree. Such a goal may be articulated as compelling the reduction of nation-state-sponsored espionage to a level that does not critically threaten national security.

Using the coercion model presented above, the policy-maker in the targeted country can more accurately identify where it is situated in the coercion dynamic. In this way, it can show whether

4 One characteristic of being at this point is that the victim had time to gather evidence of attribution from various sources. I will revisit this point later.

5 In the case of a DDOS attack, it may become clear very quickly, or it may not.

compellence measures are more appropriate than purely deterrent measures, such as establishing an already crossed red line. The next step will be to identify what actions in cyberspace the alleged offender will perceive as sufficiently threatening to favorably influence their behavior. Table 1, also taken from a previous work by the author (Hare 2010), helps to illustrate the types of malicious activities that various nation-states may consider most threatening to their national security.[6]

TABLE 1: CYBER THREATS AND TYPES OF NATION STATES

		Socio-political Cohesion	
		Weak	Strong
Power	Weak	De-stabilizing political actions in cyberspace, attacks on Internet infrastructure, criminal activities	DDOS and other major attacks on critical infrastructure [7]
	Strong	De-stabilizing political actions in cyberspace	Criminal activities in cyberspace

To get us back to the consideration of attribution, I will present two options that may be available to a country in the W-W, W-S, and S-S quadrants that is confronted with the challenge of compelling an actor associated with an S-W nation-state (bottom-left quadrant) to cease hostile cyber actions within the victim nation's sovereign territory. States in the bottom-left quadrant do not exhibit strong socio-political cohesion, and therefore perceive themselves to be vulnerable to threats to the idea of the state (legitimacy of the regime), its institution, and even its territorial integrity (Buzan 1991).[8] I will also discuss one option to be considered by a W-W, or W-S country (e.g., small European nations), when specifically confronted with a DDOS attack.

Option 1: Aggressive Engagement in International Forums

The first policy option could be a strong push via international forums for free and anonymous access to cyberspace by citizens of all nations. Actions to support this policy would entail government participation *en masse* at influential international conferences where regulatory, legal, and standards bodies debate potential measures to ease or restrict the freedom and flow of information in the domain. Adversarial countries would most likely be trying to persuade these same bodies to enact measures that would allow more state control on the flow of information. Participants from a targeted country would aggressively lobby other attendees and seek to dominate the agenda in a way that would send a clear message to the adversarial nation that actions taken at the conference are intended to counter them. This policy measure

6 This model was previously introduced for international cyber security discussions, but the perspectives are equally relevant for analyzing coercion strategies. Some of the cyber threats to national security suggested by this matrix, such as criminal actions to steal identity, may not be appropriate considerations for legitimate policy, but there are several options open to nation-states in each quadrant.

7 A distributed denial of service attack, or DDOS, occurs when many computers, usually surreptitiously controlled, are used to inundate a web server with requests and cause it to become overwhelmed to the point that service is denied.

8 For a complete discussion of this matrix and its contents see Hare, 2010.

has the advantage of being entirely executable within diplomatic channels, and being instantly "adjustable" meaning that the lobbying pressure can easily be reduced once the adversary's hostile actions have stopped. In addition to any overt actions taken, it would still be important for the victim nation to somehow communicate to the offender that these actions are in direct response to the hostile actions the victim has attributed to this adversary. For the actions to be effective as coercive measures, it must be made clear to the adversary nation that these diplomatic actions will stop once the threatening espionage and other hostile acts stop. If the adversary does not perceive that they have an opportunity to make the retaliatory actions cease, they may respond in unanticipated ways.

Option 2: Cyber Security for Dissident Organization in the Attacking Country

The second policy measure is significantly more aggressive. This option could be comprised of two related components enacted in steps. The first step would be to provide cyber security for a dissident organization countering the adversary regime. The security measures could entail providing hardware, software, and technical expertise to the dissident organization to protect their e-mail servers from the adversary's espionage and to protect the dissident organization's web presence from disruptions. The specific actions could be done in an overt manner to send a strong signal, or clandestinely to avoid causing an uncontrollable escalation in tensions. In either case, the adversary would have to be notified that the actions are being taken in response to perceived hostile acts they have sponsored. This policy action could be enhanced by communicating that if the adversary does not cease its hostile actions against the victim country's cyber assets, the victim will increase its coercive measures by conducting counter-espionage against the adversary and providing useful intelligence to the dissident organization. This second stage may be held in reserve to stress its compellence intent. However, its coercive effect can be highlighted by informing the adversary that some information has already been divulged to the dissident organization, such as information regarding the adversary's monitoring efforts of the dissident organization.[9] In any case, the adversary must be made to understand that the threat of increased counter-measures is not an empty one.

In both options, the actions would demonstrate clear connectedness and provide a potential deterrent to future cyber threats from the adversary. When a nation has demonstrated that it is willing and capable of taking action, it greatly increases the deterrent potential of the action. In all instances, it is important to signal to the target offender that these actions are taken in direct response to their hostile actions, and that the actions will cease once the offensive actions cease. The adversary may not respond at all if they don't realize that the victim nation has attributed the hostile activity to them (Libicki 2009). An unfortunate outcome of either set of measures would be that the adversary may respond to the actions in an escalatory manner. However, communicating the rationale for either option can be done in private, which would allow the offender to avoid the appearance of bending to the victim's will, an outcome that could be politically untenable for the adversary nation's regime.

Option 3: Hunker Down in the Face of a DDOS

Without any advance warning of an overwhelming DDOS attack, the victim nation will feel the effects almost immediately. There will be no period during which compellence actions like

9 This action must, of course, be balanced with the risk of exposing tradecraft.

those described above can be developed and deliberated. The unfortunate and probably the only choice will be to block originating addresses and endure the attack until the international spotlight can be turned on the likely perpetrator. According to Klimburg (2011), a senior advisor with the Austrian Institute of International Affairs, a small, Internet-dependent country can only hope to be successful with this tactic if it employs both a horizontal and vertical "whole of nation" approach to critical infrastructure protection. Such an approach would require a strong, public-private partnership that ensures a resilient and responsive infrastructure in the face of concerted attacks. Just as a country cannot predict when and where an earthquake will occur in sufficient time to evacuate all the buildings at the epicenter, a small, Internet-dependent country must make the "building strong enough" to resist attack. This option is less directly a compellence action than the first two options. In fact, the compellence theoretically does not come from resilience but from the international condemnation that would lower the attacker's international social capital to such a degree that they determine the cost to their international standing outweighs the benefits of a continued attack.[10] Therefore, the level of resilience necessary is related to the time it will take for the international community to observe and correctly characterize an event. In the case of Estonia in 2007, it took several days for them to receive support from other countries after the onset of anonymous attacks (Evron 2008).[11]

At this point, the reader may have expected to be presented with an option where a victim fights fire with fire, or a tit-for-tat response. However, several factors suggest that a DDOS counter-attack would be an ineffective response. First, it would most likely be enacted too late to be effective as a compellence measure because unless an action was preplanned, it would take time to determine which targets to strike and how to execute the attacks. Second, building extensive "botherds" of computers from which to launch an attack would be equally time-intensive, and relying on surrogate forces or criminal entities could be considered unethical options. Third, the attacking nation is most likely not as reliant as the victim on cyberspace for functions critical to national security and therefore would not be as heavily influenced by the attack. As identified in Table 1, the cyber actions most threatening to a nation in the S-W quadrant is the proliferation of information critical of the regime. As a result, a directly coercive action cannot be expected to influence the actions of the perpetrator of a DDOS attack enough to make them change their behavior. Regardless of the coercive measure taken, stronger defenses and increased resilience of the critical infrastructure must be a part of any strategy to increase the costs of conducting hostile actions in cyberspace and to help make retaliatory actions more effective.

4. ATTRIBUTION REVISITED

I now revisit the issue of attribution to determine its significance to the three policy scenarios presented above. As previous authors have done, I will address the technical and human components of attribution under each set of policy options.

10 I could write several paragraphs debating whether compellence or deterrence are even options to W-W and W-S type countries. However, the focus of this paper is on attribution, so I have entertained this scenario only to demonstrate that it conforms to the theories presented in the paper.

11 I acknowledge that it is impossible to determine if the international outrage ever did have an effect on the attackers since the attacks lasted for several more weeks. It is quite possible that they didn't cease until the attackers just got bored. This is one point that supports the previous footnote.

Attribution and Option 1: Engagement in International Forums

Aggressive lobbying for positions counter to the interests of certain regimes is not considered a hostile act under existing international conventions. Therefore, this coercive measure is not difficult to justify to internal or international audiences. There is no expectation that the victim country must explain to third parties that its action is a response to offensive actions by the adversary. Therefore, there is little need to attribute the initiating action to an international audience. However, a certain degree of human attribution is still necessary to make this option effective. The victim country must at least discover some correlation to the regime of the targeted nation-state and offer evidence that the offending actor is an entity over which the regime can exert some control if it so chooses. If the actor is one over whom a national regime cannot exert its sovereign power, then there is little chance the measure will achieve the intended coercive effect. Some evidence of this correlation may need to be conveyed to the adversary regime (potentially confidentially) to help the offender understand the link between its actions and the victim's response. Some amount of attribution may also be important when the adversary regime feigns ignorance of the event or if the adversary claims in an international forum that the aggressive diplomatic activity is an unprovoked assault. Technical attribution, even though it may have been achieved to support a determination of human attribution, is much less important in this situation. There is no policy requirement to tie an offending action to a specific machine, as the response measure is not tied to any specific machine. Ultimately, the burden is on the regime of the adversary nation to bring the hostile behavior to a halt. To do so, the regime will need to direct specific actors to alter their behavior. As long as the regime is aware of the responsible parties, knowledge of the specific networked entities used to conduct the hostile acts is of less importance.

Attribution and Option 2: Cyber Security for a Dissident Organization

This measure will clearly be considered threatening by the regime of an S-W nation. It also could be perceived by third parties as a direct challenge to the sovereignty of the targeted nation and therefore create concern in the international community. This concern will be greatest if any actions associated with the policy measures involve conducting a cyber operation within the territory of the targeted nation. In this case, there is a real possibility that the adversary nation will try to paint itself as the offended party and complain to the international community that it is experiencing an unprovoked attack on its sovereignty. Moreover, the victim nation initiating the response action must strive to demonstrate to the adversary regime that the measures are intended to be directly connected to and in response to actions determined to have been taken by the adversary.

Because of these two concerns, there is a greater requirement for some level of attribution that can be demonstrated to the adversary and, if events become public, to the international community. As with the first policy option, the focus of attribution must be the human or organization that perpetrated the offensive intrusions. Any evidence with which the victim chooses to demonstrate this linkage can be used to support attribution. It can consist of technical data, or intelligence gathered by multiple means. The argument to support attribution can also consider other events that may show a *cui bono* reason why the adversary was the most likely perpetrator of the hostile act. The duration of events may help determine attribution. An intrusion event of enduring nature may have increased the threat to national security, but it also gives the victim

with more opportunities to gather evidence of attribution using various means. Compiling several sources may allow the victim to obfuscate the specific details of any one source and avoid the unintended consequence of providing positive feedback to an adversary on how they can improve their intrusion tactics (Libicki, 2009). One of the sources could be technical attribution. Since the threatening espionage will have required two-way communications over an extended period to retrieve intelligence, it may provide more opportunities for a victim to identify valid source codes and overcome one of the largest barriers to attribution identified by Clark and Landau (2010)—the multi-stage attack. The espionage-motivated intrusions will most likely be against several information system targets. If the victim unravels the intrusion events, the complexity may provide an aggregation of evidence from several systems (Lewis 2011). Although the challenges of attribution do not change in this situation, the opportunities to achieve it may increase.

Attribution and Option 3: Hunker Down in the Face of a DDOS

As in option 1, enacting a strong defense that ensures the availability of services in the face of a DDOS attack and the regeneration of data in the case of a server crash would not be considered hostile acts. These are purely actions of critical infrastructure protection. As such, the only attribution necessary is that required to block offending intellectual property addresses at the target. In most cases, this technical attribution back to the last hop is easy to achieve. Unlike a case of espionage or more surgical cyber exploitation of critical infrastructure, there would be less exposure of state secrets if the event were publicized. In that case, it would be less politically risky to invite the broader community to advise in the defense of the nation. Therefore, cyber security experts from around the world could be invited to participate in the defensive actions and to help build a clearer picture of attribution. The necessary level of attribution is contingent on the case to be made to the international community. As in the previous case, the argument to support attribution can also consider other events that may show a *cui bono* reason why the adversary was the most likely perpetrator of the hostile act. For example, political disputes between Estonia and a neighbor nation provided the Estonian responders with a clear suspect. Also, the duration and level of a DDOS attack may help the defenders compile a sufficient level of human attribution based on the cumulative technical attribution. In the attacks on Estonia, on-line forums in certain communities were abuzz with discussions and instructions on how to participate in the attacks, which contributed to the determination of human attribution (Herzog 2011). Whether or not a particular regime is directly engaged in an activity, the international community may still consider it culpable and responsible, either legally or politically, for influencing the malicious behavior of actors under its influence. If a regime is unwilling to deal with the malicious actors or claims it is unable to do so, it could cause that regime to lose political capital. However, the policy-maker in the victim country must be willing to accept the fact that the issue of attribution is irrelevant if a loss of stature in the international community is irrelevant to the probable perpetrator.

5. CONCLUSIONS

There is no question that the anonymity and ease of international interaction in cyberspace increases opportunities for malicious activity. However, the coercion challenge is no more difficult in cyberspace than in other domains. The amount of evidence required to support an attribution argument will depend on the political situation at the time of the response action and the adversary's receptiveness to the victim's efforts to link the two actions. If there are many other factors pointing to a specific adversary as the likely instigator (the *cui bono* test), then that adversary will most likely be the focus of attribution and the international community will more readily support a claim of attribution without specific evidence.

In this paper, I have argued that unequivocal attribution is not required to enact a retaliatory measure and that attribution may be determined only after the measure is enacted successfully. However, for the compellence measure to be successful, the adversary must know that the victim has attributed the hostile actions to them, that the compellence measure is in retaliation for the offending action, and that the pain will stop only after the adversary has complied with the victim's demands. Confident assessment of human attribution will strengthen the effect of coercive responses. While the anonymity provided by cyberspace allowed the offender to conduct a threatening act that is not visible to others, it also enables a flexible coercion strategy. For example, it allows the communication and application of the compellence measure to be conducted privately. The benefit here is that the victim can plan its response actions with less concern about the influence of third parties or the demands of conclusive attribution.

For the three potential policy options discussed in this paper, attribution is a useful but not a required component of a coercion strategy. At the international level, national decisions are based on political considerations over legal ones. In a legal situation, attribution may be a requirement, but in a case of political calculus, attribution is one factor that must be balanced with all other political considerations of national security.

ACKNOWLEDGEMENTS

I would like to thank Col. David Fahrenkrug for giving me the idea of changing the coercion discussion from deterrence to compellence (and lending me the books to learn what that means). I would also like to thank my colleague Jeff Goldman for insightful comments on the first draft of my paper and the assigned reviewers for having forced a more international perspective in my presentation.

REFERENCES

Alperovitch, D. 2011. "Towards Establishment of Cyberspace Deterrence Strategy." In 3rd International Conference on Cyber Conflict (ICCC), 1–8. Institute of Electrical and Electronics Engineers.

Australian Government. "Mitigating the Cyber Threat". Government Document. Connecting with Confidence: Optimising Australia's Digital Future. http://cyberwhitepaper.dpmc.gov.au/white-paper/security-and-resilience-in-the-online-environment/mitigating_the_cyber_threat.

Boebert, W. Earl. 2010. “A Survey of Challenges in Attribution.” In Proceedings of a Workshop on Deterring Cyberattacks:Informing Strategies and Developing Options for U.S. Policy, 41–52. Washington, D.C.: The National Academies Press.

Buzan, Barry. 1991. People, States, and Fear: The National Security Problem in International Relations. 2nd ed. Boulder: Lynne Rienner.

Byman, Daniel, Matthew C. Waxman, and Eric Victor Larson. 1999. Air Power as a Coercive Instrument. Santa Monica, Ca: Rand Corporation.

Clark, David, and Susan Landau. 2010. “Untangling Attribution.” In Proceedings of a Workshop on Deterring Cyberattacks:Informing Strategies and Developing Options for U.S. Policy, 25–40. Washington, D.C.: The National Academies Press.

Deibert, Ronald, and Rafal Rohozinski. 2009. Tracking Ghostnet. Intrusion Analysis. Information Warfare Monitor. Toronto: Centre for International Studies, University of Toronto. http://www.scribd.com/doc/13731776/Tracking-GhostNet-Investigating-a-Cyber-Espionage-Network.

Evron, Gadi. 2008. “Battling Botnets and Online Mobs.” Georgetown Journal of International Affairs 9: 121.

Gorman, Siobhan. 2009. “Electricity Grid in U.S. Penetrated By Spies.” Wall Street Journal, April 8, on-line edition, sec. Technology.

Hare, Forrest. 2010. “The Cyber Threat to National Security: Why Can’t We Agree?” In Conference on Cyber Conflict Proceedings 2010, 211–226. Tallinn, Estonia: CCD COE Publications.

Hare, Forrest. 2011. “The Interdependent Nature of National Cyber Security: Motivating Private Action for a Public Good”. Doctoral Dissertation, Fairfax, Va: George Mason. http://u2.gmu.edu:8080/handle/1920/6312.

Harknett, Richard. 1996. “Information Warfare and Deterrence.” Parameters (Autumn 2006): 93–107.

Herzog, Stephen. 2011. “Revisiting the Estonian Cyber Attacks: Digital Threats and Multinational Responses.” Journal of Strategic Security 4 (2) (July 1). doi:10.5038/1944-0472.4.2.3. http://scholarcommons.usf.edu/jss/vol4/iss2/4.

Klimburg, Alexander. 2011. “Cyber Security Und Schutz Kritischer Infrastrukturen”. Newsletter. Newsletter der GIT Gesellschaft für Informations- und Kommunikationstechnik im OVE. http://git.ove.at/newsletter/GIT_Newsletter_05_2011.htm#klimburg.

Landler, Mark, and John Markoff. 2007. “Digital Fears Emerge After Data Siege in Estonia.” The New York Times, May 29, sec. Technology. http://www.nytimes.com/2007/05/29/technology/29estonia.html.

Lewis, James. 2011. “Rethinking Cybersecurity – A Comprehensive Approach” presented at the Sasakawa Peace Foundation, September 12, Tokyo. http://csis.org/publication/rethinking-cybersecurity-comprehensive-approach?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+CSIS-Cybersecurity-Related-Publication+%28Cybersecurity+-+Related+Publication%29.

Libicki, Martin C. 2009. Cyberdeterrence and Cyberwar. Santa Monica, Ca: RAND Corporation.

Lukasik, Stephen. 2010. “A Framework for Thinking About Cyber Conflict and Cyber Deterrence with Possible Declaratory Policies for These Domains.” In Proceedings of a Workshop on Deterring Cyberattacks:Informing Strategies and Developing Options for U.S. Policy, 99–121. Washington, D.C.: The National Academies Press.

Schelling, Thomas C. 1966. Arms and Influence. Yale University Press.

Shanker, Thom, and Elisabeth Bumiller. 2011. “Hackers Gained Access to Important Files, Pentagon Says.” The New York Times, July 14, sec. World. http://www.nytimes.com/2011/07/15/world/15cyber.html.

Taipale, K. A. 2010. “Cyber-Deterrence.” SSRN eLibrary (April). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1336045.