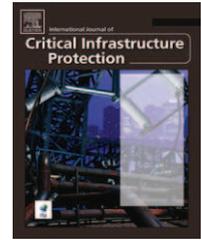


Available online at www.sciencedirect.com
SciVerse ScienceDirect
journal homepage: www.elsevier.com/locate/ijcip

A signaling framework to deter aggression in cyberspace

Mason Rice^a, Jonathan Butts^b, Sujeet Shenoi^{a,*}

^a Department of Computer Science, University of Tulsa, Tulsa, Oklahoma 74104, USA

^b Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio 45433, USA

ARTICLE INFO

Article history:

Received 24 September 2010

Accepted 11 February 2011

Published online 11 March 2011

Keywords:

Cyber operations

Signaling strategies

Deterrence

ABSTRACT

During the Cold War, the United States and the Soviet Union constantly maneuvered to achieve superiority. When one nation was perceived to overstep its bounds, the other would signal its discontent by moving aircraft carrier groups, conducting military exercises, pursuing diplomatic actions or enforcing embargoes. These clear, but nuanced, signals may well have averted nuclear exchanges.

The speed of the Internet coupled with its global connectivity and inextricable links to critical infrastructure assets render signaling just as important in cyberspace, especially as nation states and other actors are investing in cyber operations capabilities. This paper presents a flexible and intuitive framework for adversary–defender interactions involving ensembles of adversary stimuli and defender signals. Scenarios involving cyber operations on the electric power grid are used to clarify the signaling goals and corresponding “plays” executed by a defender in response to adversary actions.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The 1972 Anti-Ballistic Missile (ABM) treaty between the United States and the Soviet Union prohibited the development and testing of ABM systems. However, soon after the treaty was ratified, the US detected Soviet “cheating” via a highly classified feature of Project MELODY that intercepted Soviet missile tracking radar signals [1]. During subsequent negotiations in Geneva, Secretary of State Henry Kissinger looked his Soviet counterpart in the eye and revealed the dates and times when the Soviets cheated on the treaty. The cheating stopped and the Soviets began a “mole hunt” for the spy who gave the information to the United States. Kissinger sent a clear signal to the Soviet Union and America got its way without compromising its MELODY sensors.

Signaling is a highly nuanced mode of communication that is used primarily in the animal kingdom. Guided by human analysis and introspection, signaling has been used very effectively in the geopolitical realm to deter aggression.

Signaling actions ranging from cat-and-mouse submarine patrols to elevated DEFCON levels kept the Cold War from escalating. Saddam Hussein may well be alive had he not misread US signals before Gulf War I and again in the months before Gulf War II. But in no other battlespace may signaling be as important as in the global Internet environment.

Because of its inextricable links with the critical infrastructure, the Internet is vital to the security of nations and the well-being of citizenry. Attacks during World War II targeted strategic infrastructures; cities were fair game — London, Dresden and, ultimately, Hiroshima and Nagasaki. Internet attacks may not kill millions like nuclear weapons, but sustained, large-scale attacks could be devastating. How would Americans cope if much of the electric power grid were to go down – and stay down – for six months? Such a long-term outage would result in mass human migration; populations in major cities could drop to pre-1850 levels.

* Corresponding author.

E-mail address: sujeet@utulsa.edu (S. Shenoi).

Nation states and other actors employ cyber operations to gain economic, strategic and other advantages [2]. Cyber operations involve the attack, defense and exploitation of electronic data, knowledge and communications, possibly impacting infrastructure assets and human life [3]. It is, therefore, vital to develop flexible signaling strategies that can deter aggression in the global Internet environment.

This paper describes a general signaling framework that is derived from strategic (e.g., diplomatic and military) signaling techniques. One example of signaling involves giving the adversary the appearance that the defender is either unaware of the adversary's activity or that the activity was detected by chance. Another example is reflexive signaling, which is designed to appear as an immediate reaction to some stimulus. The principal signaling constructs, along with their themes and variations, are discussed using several scenarios involving cyber operations on the electric power grid. The power grid provides a rich environment for clarifying the principal issues related to signaling. Also, it is a very relevant case study because some nation states are reportedly conducting cyber operations on the US power grid [4,5].

2. Cyber operations and signaling

Owens, et al. [6] argue that the “seductive” quality of cyber operations may well increase the likelihood of their use. Much like playing a video game, a cyber operation is clinical in nature and is often executed remotely and potentially anonymously. Also, they are seemingly non-lethal — like tasers. According to one study [7], while the number of fatalities due to police action decreased when police were armed with tasers, the number of instances involving the use of force increased dramatically because police were more willing to use the non-lethal tasers. Indeed, before tasers, the police often used friendly persuasion or found some other way to resolve the matter without the use of force.

Cyber operations have other characteristics that promote their use. Attack and exploitation tools are inexpensive to build and deploy, and they are highly replicable. Unlike traditional military maneuvers, cyber operations are conducted in seconds. Also, cyber operations are difficult to detect and attribute. Attackers can mask themselves and their exploits, and disappear into the Internet cloud.

Signaling in cyberspace requires a nuanced approach because of the shadowy nature of adversaries, and the ambiguities related to their capabilities, intentions and targets. To be effective, signaling in cyberspace must be clear, fast and sophisticated. Also, the signaling entity often has to preserve the secrecy of the detection mechanisms and be cognizant that signals propagate beyond their intended targets because of Internet connectivity.

3. Adversary and defender interactions

Signaling involves interactions between an adversary and a defender that are spread over space and time. A typical Cold War example involved the detection of Soviet submarine activity near US territorial waters [8]. To signal its discomfort,

the US moved strategic bombers to a higher state of readiness, knowing that Soviet satellites would report the bomber activity. Because the responsive signal was proximate in time and proportionate in scale, the US was (rightly) confident that the Soviets would correctly interpret the action as a response to their initial submarine activity and would not see it as an unrelated event or an escalation. The clear American signal and the associated counterthreat forced the Soviet submarine to retreat.

Fig. 1 provides a generic representation of the interactions between an adversary and a defender from the perspective of the defender. The adversary and the defender have actuators and sensors that are separated by a notional barrier or membrane. Actuators are of two types — stimulus actuators that produce adversary actions and signal actuators that produce defender signals. Sensors deployed by the defender detect adversary stimuli while those deployed by the adversary detect defender signals. The defender has an analysis component that processes sensor information and determines and initiates the appropriate signals. The analysis component also enables the defender to perceive the state (of mind) of the adversary when producing a stimulus and the (possibly different) state of the adversary after receiving the signal.

In general, adversary and defender interactions involve ensembles of stimuli and signals over space and time. We assume that each stimulus and signal occurs at a unique instant of time. Also, it is not necessary for stimuli and signals to alternate. Furthermore, the interactions could begin with an attacker stimulus or a defender signal.

The framework is not limited to modeling interactions involving a single adversary and a single defender. Scenarios involving multiple independent or cooperating adversaries and/or defenders can be modeled using a single diagram as in Fig. 1. However, scenarios involving multiple independent defenders would require multiple diagrams.

4. Actuators and sensors

Actuators are symbolic constructs that produce benign actions or malevolent actions. Benign actions, such as passive surveillance and tagging (e.g., a Post-it note stating “Kilroy was here!”), cause no specific damage to assets aside from psychological effects. Malevolent operations, which involve potentially harmful actions, include active probing, exfiltration, system manipulation, malware installation and denial of service.

In general, adversaries and defenders can execute benign and malevolent actions in cyberspace as well as in other realms (e.g., diplomatic, information, military and economic domains). Interactions involving benign and/or malevolent actions in these domains are readily modeled using our adversary–defender framework. However, since our focus is on cyber operations, we assume that the adversary's actions are limited to cyberspace, i.e., the stimulus actuators are only used by the adversary to conduct cyber operations. On the other hand, the defender may employ signal actuators to perform actions in cyberspace and in other domains.

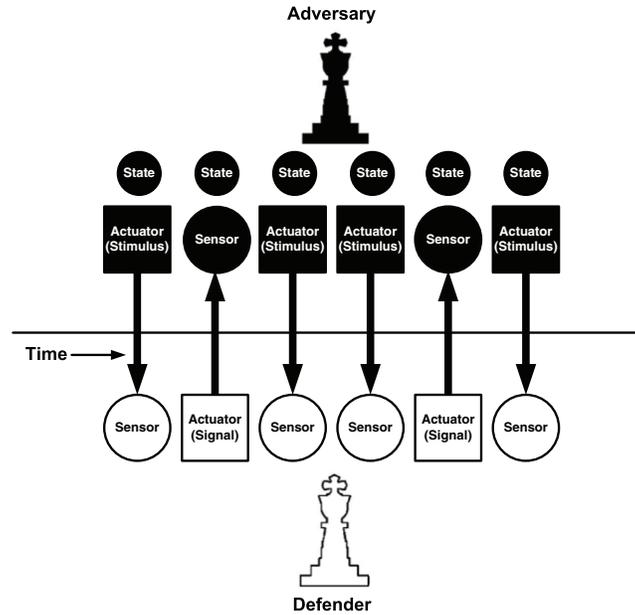


Fig. 1 – Adversary and defender interactions.

Sensors are used by the defender to detect stimulus actions and by the adversary to detect signal actions. Sensor attributes include modality, location and range, sensitivity, credibility and secrecy. The modality of a sensor refers to its detection mechanism (e.g., electronic, thermal, magnetic, radiant and chemical) [9]. The location and range of a sensor specify the space in which the sensor can operate effectively. Sensitivity refers to the ability of a sensor to detect stimuli and signals; cyberspace sensors may be tuned to detect specific viruses and worms, rootkits and network probes. The credibility of a sensor is a function of its reliability and durability; reliability refers to the ability to correctly classify stimuli and signals while durability refers to the ruggedness of the sensor and its tamper resistance.

The secrecy of a sensor is an important attribute in our discussion of signaling. The attributes of a sensor determine its secrecy. In general, if one attribute of a sensor is classified, the existence and/or use of the sensor may be classified. However, the existence of a sensor may be public knowledge, but its attributes could be classified. For example, the location and modality of the US underwater sound surveillance system (SOSUS) may be known, but its sensitivity is a closely guarded secret [10].

5. Signaling goals

The adversary's decision to conduct an operation involves three primary variables: (i) perception of the benefits of a stimulus; (ii) perception of the costs of the stimulus; and (iii) perception of the consequences of inaction [11]. The perceived benefits and costs of a stimulus (including inaction) have relative values to an adversary and associated probabilities that feature in the adversary's decision calculus. This section describes the signaling goals on the part of the

defender in response to adversary stimuli (including null stimuli).

The defender has three basic ways to deter an adversary. The first is to credibly threaten and/or deny the adversary the benefits or gains sought [11]. The second is to credibly threaten and/or impose severe costs on the adversary. The third is to encourage restraint by convincing the adversary that inaction is the best possible outcome. In general, the defender may select one or more of these options to deter the adversary.

Denying benefits by the defender involves defensive and offensive capabilities and activities [11]. For example, an anti-ballistic missile system that intercepts adversary missiles is an example of an operational capability that provides deterrence by credibly threatening to deny future benefits.

In circumstances marked by a pronounced asymmetry of stakes and confrontation with a risk-acceptant adversary, denying benefits takes on increased importance [11]. Such adversaries tend to discount the severity and/or the likelihood of the costs that a defender might impose. An example nation-state actor is North Korea, which has sophisticated cyber operations capabilities but little domestic reliance on cyberspace [12].

Deterrence by cost imposition involves convincing the adversary that the costs incurred as a result of the adversary's planned stimulus are severe and highly likely [11]. Cost imposition includes all the domains of power. The key challenge to improving the effectiveness of deterrence by cost imposition is to overcome the adversary's perception that it can deter a counterattack or that (for political or other reasons) the defender will simply choose not to counterattack. Tit-for-tat actions are often used in the intelligence realm. When a sensitive government system is probed by an adversary, the defender may choose to launch a comparable probe on an equivalent asset belonging to the adversary.

Encouraging adversary restraint can be accomplished in two ways [11]. First, the defender can signal the adversary about the benefits of continued restraint. Second, the defender can take actions that mitigate the costs of restraint perceived by the adversary. For example, the defender's doctrine might call for cyber operations to be conducted in a manner that would inadvertently mislead the adversary about the nature of the defender's objectives, or might impose unintended and unnecessary costs on the adversary. Either of these circumstances could result in the adversary choosing to escalate a conflict that would otherwise be limited. Therefore, it is crucial that signaling actions are clearly communicated to and understood by the adversary.

In summary, the defender's signals must convince the adversary that its stimuli will: (i) fail to achieve their objectives and reap the benefits sought, (ii) incur severe costs to the adversary that would outweigh the perceived benefits, and/or (iii) cause the adversary to suffer an outcome that would be worse than if it had pursued no action [11].

6. Signaling constructs

The general signaling constructs described in this section are derived from strategic signaling techniques.

6.1. Primitive signals

Primitive signals are used in the adversary–defender interaction framework individually or collectively to create complex, nuanced signal ensembles. The two types of primitive signals are null signals and simple signals.

6.1.1. Null signals

A null signal involves no signaling action on the part of the defender upon receipt of a stimulus from the adversary. The decision to tolerate the stimulus could be driven by a desire to conduct additional surveillance, to maintain the secrecy of the sensor or because the stimulus does not exceed a threshold. A Cold War example of toleration involved the use of US “gatekeeper” submarines off the Soviet ports of Petropavlovsk and Vladivostok, and near the Kola Peninsula for the express purpose of collecting data about Soviet nuclear submarines [13].

6.1.2. Simple signals

A simple signal involves a signaling action by the defender either unilaterally or in response to a stimulus from the adversary. As mentioned above, the defender may send the signal in cyberspace or some other (e.g., diplomatic, information, military or economic) domain. The signal may express attitude or emotion (e.g., displeasure), capability (e.g., show of force), knowledge (e.g., awareness of the stimulus), intent (e.g., retaliation or resolve), presence (e.g., location) and/or personality (e.g., friendliness or hostility).

The signaling action itself can be broadly categorized as an emblem, illustrator, regulator, adaptor or affect display [14]. An emblem is a movement or act that is a substitute for words (e.g., shaking a fist or waving as a greeting). An illustrator accompanies, modifies or exemplifies a communication

(e.g., pointing action). A regulator is a movement that maintains or changes the communicative role (e.g., nodding to convey agreement or waving an arm to express dissent). An adaptor is related to an emotional state (e.g., the protective movement of folding the arms across the chest). An affect display is primarily related to facial expressions, but it does not take much imagination to envision how a defender can employ such an action in cyberspace or some other domain in conjunction with its rhetoric.

6.2. Signaling plays

Signaling plays are composed of primitive signals. The plays can be offensive, defensive, combined offensive–defensive or neutral. This section describes simple signaling plays and ensemble signaling plays, which are sequences of primitive signals devised by the defender to convey a nuanced message to the adversary.

6.2.1. Simple signaling plays

Simple signaling plays are composed of a single primitive signal (i.e., null signal or simple signal). An example of a null signaling play in cyberspace involves a defender finding a Trojan horse planted by an adversary, but choosing not to act because of an ongoing espionage investigation. Another example is an adversary exfoliating classified information about a weapons system, but the defender opts for a null signal because the information is part of a canard or setup.

An example of a simple signaling play in cyberspace is to block network access from a specific set of IP addresses from where an attack has been launched. At a minimum, this signaling play would indicate the defender's awareness and displeasure. Another example is the execution of a denial-of-service attack on the adversary's assets in response to a cyber operation. The counterattack would indicate detection capability, displeasure, hostility and resolve on the part of the defender.

Two useful signaling plays involve the use of reflexive signals and random signals.

Reflexive signaling play. A reflexive signaling play is intended to be perceived as strictly reactive by the adversary, similar to the patellar reflex. A Cold War example is “launch on warning”, in which the US doctrine was to launch its strategic nuclear arsenal simply upon detection of an impending Soviet attack. Launch on warning requires knowledge of the characteristics of an attack and unimpeachable command and control procedures.

The cyberspace equivalent of launch on warning involves the defender disconnecting itself from the external Internet if crippling cyber operations from a sophisticated adversary are imminent. In fact, legislation has been proposed that would grant the US President the ability to declare a national cyber emergency, which would require service providers and search engine companies to sever their external connections [15].

A reflexive signal is designed to appear as an immediate response to a benign or malevolent operation. The specific signal may be determined in advance based on the attributes of the stimulus (e.g., originator, type and location).

A reflexive response may not necessarily involve memory. By limiting memory in a reflexive response, the defender can signal and then “forgive and forget” or ensure that the reflexive action remains consistent. Note, however, that a reflexive action may be adjusted as priorities and conditions change.

Technologies are under development to implement reflexive signaling in cyberspace. For example, the network-centric collaborative targeting (NCCT) system [16], which determines the location of a target with minimal human intervention using a network of sensors, could be leveraged to perform reflexive signaling.

Reflexive actions can be purely defensive in nature. One example is the Homeland Security Advisory System (National Threat Advisory) with its five color-coded categories ranging from “low”(green) to “severe”(red). The threat levels change as different stimuli are detected. Various actions are prescribed at each threat level. For example, actions taken during a “high” (orange) condition include coordinating security efforts with law enforcement agencies, national guard and the military, taking additional precautions at public events, preparing to execute contingency procedures, and restricting access to threatened facilities. Other national warning systems with a reflexive signaling component are the Department of Defense’s Defensive Condition (DEFCON) and Information Condition (INFOCON).

Random signaling play. A random signaling play may be used to confuse the adversary. Such a play can facilitate other operations undertaken by the defender while appearing to be random. If the adversary detects an action (e.g., re-routing network traffic or conducting a security audit), then the adversary must determine if the action is a signal that a cyber operation was detected by the defender or if the action is an unrelated (previously scheduled) event. Note that designed random signals are proactive in nature, whereas many simple signals are reactive.

Examples of simple signaling plays. Table 1 presents examples of simple signaling plays, including reflexive and random signaling plays. Note that the signaling plays are categorized into four groups based on their intent: offensive, defensive, offensive–defensive and neutral.

6.2.2. Ensemble signaling plays

An ensemble signaling play is a sequence of primitive signals devised by the defender to convey a nuanced message to the adversary in response to one or more stimuli. Indeed, an ensemble signaling play is the defender’s portion of a conversation or, possibly, a game of strategy intended to inform, entertain or persuade the adversary. In general, the signals in an ensemble are designed by the defender to respond to adversary stimuli taking into account the defender’s perception of the state of the adversary (Fig. 1).

A classic ensemble signaling play is “shepherding”. Shepherding involves the orchestration of signals to subtly guide the actions of the adversary. A classic Cold War example is the CIA’s use of the PALLADIUM system during the Cuban Missile Crisis. PALLADIUM was designed to deceive radar systems into seeing and tracking ghost aircraft [1]. In one instance, PALLADIUM was used to create ghost aircraft that

were always just ahead of pursuing Cuban fighters, effectively shepherding the fighters away from a sensitive area.

In cyberspace, shepherding can be conducted very effectively using honeypots and honeynets. Honeypots are traps to detect and/or deflect unauthorized access to computer systems and networks. A honeynet is a high-interaction honeypot environment with systems, applications and services [17]. A honeypot is typically static in nature, while a honeynet appears as a live network to an attacker. In both cases, however, the adversary believes it is conducting operations on a genuine system.

Honeypots and honeynets can be designed to draw attacks away from real assets. When an adversary penetrates a sensitive system or network, an ensemble signal could be used to draw it to decoy assets in a honeypot or honeynet. Upon entering the decoy system, the adversary is monitored extensively and valuable information is collected about its tactics, techniques and tools.

Another shepherding strategy involves the defender executing a series of seemingly random cyber operations on adversary assets upon detecting a stimulus. In this case, the defender’s intent is to distract and redirect the adversary, creating a cat-and-mouse situation.

6.3. Signaling contexts

Signaling plays comprise a simple signal or multiple simple signals that can be categorized as offensive, defensive, combined offensive–defensive or neutral. The play that the defender implements must align with the proper context (e.g., conflict resolution or territorial defense) based on the state of the adversary.

Note that the meaning of a signal to the adversary could vary widely depending on the context. For example, suppose the defender performs a port scan on the adversary’s system. If the adversary and defender have had little or no previous interaction, the scan could be a test or a friendly gesture that points to a vulnerable firewall. However, if the adversary and defender have tense relations, the port scan could be construed as a warning that the adversary is trespassing on the defender’s network.

Signaling contexts are well established in animal communication. An animal may have a limited signaling repertoire, but each signal may have a different meaning depending on the context in which it is used (e.g., conflict resolution, territorial defense, environment and autocommunication) [18]. In the context of conflict resolution, signals are likely to indicate intentions, levels of commitment and offensive capabilities. Territorial defense, which initially involves conflict resolution, is associated with maintenance and safeguarding a particular location and demarcating boundaries. Signals in the environmental context are used to provide information about conditions external to the defender and/or adversary. Autocommunication is used to identify the differences between the emitted and received versions of a signal; this is often used to determine the ambient conditions in the environment.

A variety of signaling plays can be constructed for a given scenario. Just like in animal communication, there are constraints in the physical and cyber environments that limit

Table 1 – Examples of simple signaling plays.

Offensive	<p>Null signal: Show goodwill by not attacking; conduct secret invasive surveillance.</p> <p>Simple signal (reflexive): Launch an attack when an imminent threat is detected.</p> <p>Simple signal (random): Sever communication links to degrade the adversary's ability to communicate while giving the appearance that the cause was accidental.</p> <p>Simple signal (other): Actively probe the adversary's assets; launch a tit-for-tat and/or mirror image attack; deny service; disrupt the adversary's operations; destroy the adversary's data.</p>
Defensive	<p>Null signal: Show goodwill or ignorance by not assuming a defensive posture; conduct passive surveillance; conduct secret active surveillance; sacrifice a less important system in an effort to study the adversary's attack methods.</p> <p>Simple signal (reflexive): Sever Internet connections when an attack is imminent or underway; change the National Threat Advisory status and/or INFOCON status.</p> <p>Simple signal (random): Deploy blue teams to identify and eliminate vulnerabilities; deploy open sensors; re-route traffic.</p> <p>Simple signal (other): Announce the deployment of open and secret sensors.</p>
Offensive–Defensive	<p>Null signal: Display obliviousness or goodwill by not acting.</p> <p>Simple signal (reflexive): Change the DEFCON status.</p> <p>Simple signal (random): Announce that cyber operations forces are spread throughout the world and attacks may not be launched from within the geographical boundaries of the defender; threaten severe penalties to an adversary who conducts cyber operations on the defender; conduct a show of force to display capabilities; conduct a random security audit.</p> <p>Simple signal (other): Threaten an adversary with military and/or economic force; offer incentives for restraint; bluff an adversary with capabilities that are not yet weaponized.</p>
Neutral	<p>Null signal: Maintain the status quo by not acting.</p> <p>Simple signal (reflexive): "Growl" by actively pinging border routers worldwide.</p> <p>Simple signal (random): Create a mystery (e.g., slow communication links or drop a large number of packets); conduct a show of force.</p> <p>Simple signal (other): Launch an attack on oneself using a known adversary capability; signal the discovery of an event that did not occur; offer assistance to the adversary (e.g., blue team services); perform benign tagging; send friendly alert messages by pinging the adversary's assets.</p>

the ability to signal. In animal communication, the process of finding the best signal is called optimization [18].

In general, a defender will face adversaries whose political, cultural, ideological, religious and idiosyncratic values vary considerably [11]. These differences complicate and influence the adversary's perceptions of the defender's signals. Therefore, care must be taken to select and monitor a signaling play to ensure that it is not misinterpreted (or unnoticed) by the adversary. The defender must also consider the potential for miscalculation and select a play that is optimized for the context and that will convey the appropriate message.

7. Signaling challenges and pitfalls

Signaling can be used to demonstrate situational awareness, effective command and control, forward presence, integration and interoperability of sensors and signal actuators, active and passive defenses and global operational capability. However, certain challenges and pitfalls can hinder effective signaling, in particular, attribution, unintentional signals and escalation.

7.1. Attribution

Attribution in cyberspace is a major challenge. However, there are at least three factors that may facilitate attribution [19]. First, for a variety of reasons, an adversary may choose to reveal to the defender that it is responsible for a cyber

operation. Second, certain cyber operations might share technical features that convey an identifiable "signature". Third, the defender may have out-of-band information that points to the adversary, such as information from a spy in the adversary's command structure or high-quality signals intelligence.

Even if the attacker is not identified, it might be possible to hold some entity – such as a nation state that has jurisdiction – responsible for stopping the attack and identifying the attacker [12]. While attribution is a challenging and often indeterminable problem, signaling is still effective because a defender can always send signals to multiple adversaries.

7.2. Unintentional signals

Certain actions taken by the defender are not intended to be signals, but may be construed as signals by the adversary [20]. Research has shown that potentially dangerous developments in past crises occurred because civilian authorities did not thoroughly understand the military operations they were contemplating [20]. An example is the global nuclear alert that occurred in 1960 as a result of a vague request by US Secretary of Defense Thomas Gates to the Joint Chiefs of Staff. Secretary Gates' request came from Paris, where Eisenhower and Khrushchev were attending a summit. Tension over the shootdown of a U-2 plane in Soviet airspace two weeks earlier had already undermined the summit and the provocative alert dealt a fatal blow to the summit. Gates later testified before Congress that he had only meant to test the military alert system.

In the cyberspace environment, random incidents can lead to unintentional signals (e.g., hardware failures, software flaws and operator errors). Leaders and other decision makers who may not fully understand the context and the adversary's state of mind may send the wrong signal. Like the military alert ordered in 1960 by Defense Secretary Gates, a cyber alert – such as an INFOCON status change for training purposes – in a tense geopolitical environment could be misinterpreted by the adversary as a cover for defensive preparations as a prelude to full-scale cyber operations.

7.3. Escalation

Signaling can be very useful to express discontent and hostility. Military signals (alerts) enable both the defender and adversary to convey concern and determination, effectively supplementing verbal diplomacy [20]. The signals could be positive or negative depending on numerous factors, the most important of which is mutual perception. Even defensive alerts are prone to misinterpretation. An alert on one side increases the risk of provoking a reciprocal alert, which could result in a vicious cycle of escalating alerts and actions.

A fundamental issue in crisis management is to formulate a policy that strikes a reasonable balance between the need to establish a credible threat and the need to demonstrate nonaggression to the adversary [20]. The weights attached to these objectives vary according to the circumstances, with some interactions needing to show resolve while others attempting to allay fears on the part of the adversary.

A tit-for-tat action can be a clear non-escalating signal. A Cold War example occurred when the US Embassy was told by the Soviet leadership that the entire country outside Moscow was closed to travel by American diplomats [21]. In response, the State Department instituted similar restrictions on Soviet diplomats in Washington just before Ambassador Dobrynin's speaking engagement in Chicago. The Soviets got the point and lifted the travel restrictions; the State Department reciprocated almost immediately.

A cyberspace example involves the discovery that the adversary has planted malware in the defender's networks. In response, the defender may consider executing attacks against the adversary, which could escalate the actions on both sides. It might be more prudent for the defender to signal its awareness and displeasure, but this may not always be the optimal signal in the particular context.

In other cases, it may be necessary for the defender to send a strong signal to force the adversary to cease its cyber operations and ultimately stop any escalation. This could occur, for example, when the adversary is launching large-scale denial-of-service attacks on the defender's telecommunications networks. The defender may opt to respond with attacks that target the cyber assets, physical facilities and personnel associated with the denial-of-service attacks.

8. Signaling plays in the electric power grid

Numerous signaling plays can be constructed based on the adversary's stimuli and state and the defender's signaling

actions. The plays are simply ensemble signals that are created by interleaving adversary stimuli and primitive signals on the part of the defender. As mentioned above, signaling plays can be categorized as: offensive, defensive, combined offensive–defensive and neutral. These plays can be used to express attitude or emotion, capability, knowledge, intent, presence or personality, or various combinations thereof. This section describes signaling plays corresponding to three scenarios involving cyber operations on the electric power grid.

8.1. Null signal scenario

A federal government security expert is embedded as an employee in the control center of a privately owned power generation facility, which provides electricity to critical military and intelligence agency installations. Only the CEO of the company knows that the federal security expert is an embedded employee.

During the course of his work, the security expert detects – using a secret method – a fake administrator account on a network device that controls VPN tunneling to the control center. The parent government agency determines that the fake administrator account was planted by a nation-state adversary. To protect the secrecy of the embedded government employee and the detection method, a decision is made to remain silent and tolerate the intrusion in an attempt to study the tactics, techniques and tools of the adversary. Also, a decision is made to monitor the fake account for malevolent activity.

8.2. Ensemble signal scenario

This scenario builds on the null signal scenario. In this case, a decision is made by the government agency to deter the adversary by denying benefits and imposing costs, but in a way that allows the defender to learn the tactics, techniques and tools without compromising the secrecy of the embedded employee and detection method. Otherwise, the fake account created by the adversary could simply be removed.

To achieve its ends, the defender creates a honeynet that appears to contain several fault control sensors. The entrance to the honeynet is through the network device that contains the fake account. An initial random (simple) signal is sent by creating a file in the shared operator workspace that announces the installation of the fault control sensors and that information about the sensors is stored with configuration management data in certain files in the honeynet.

Upon entering the honeynet, the adversary believes that it can manipulate the fault control sensors on the power grid and tests this ability, which triggers secret sensors in the honeynet. In response, the defender signals annoyance by briefly flooding the adversary's communication link. This "emblem" signal indicates to the adversary that the defender is aware of the intrusion and can slow, if not stop, further network intrusions.

However, the adversary is not deterred by the emblem and continues to conduct cyber operations on assets in the honeynet. In response, the defender sends two signals. The first signal is an emblem that conveys the defender's awareness of the stimulus; this emblem signal takes the

form of an email to the adversary indicating the exact time of each manipulation of the fault control sensors (like Kissinger's message to his Soviet counterpart). The second signal is a denial-of-service attack on the machines in the network segment used by the adversary to conduct its cyber operations. This signal, which is intended to demonstrate the defender's resolve and hostility, serves as a regulator (i.e., the defender assumes the speaking role in the conversation) and as an illustrator (i.e., the defender indicates the location of the adversary's attacking machine).

The defender could have chosen to plant information on one of the attacking machines to indicate that it was tipped off by a mole in the adversary's organization. Alternatively, the defender could have credited a third party with discovering the adversary's cyber operations. This was likely the case in 2005 when the Bush administration disclosed that it was working with other nations to intercept weapons and missile systems bound for Iran, North Korea and Syria [22]. In particular, senior Bush administration officials stated that Pakistan was "helpful" in tracking down parts of the global nuclear network. By naming Pakistan as the source of the information, the US concealed the use of secret sensors it may have employed. Thus, misleading and masking actions were used to protect US detection methods.

8.3. Reflexive signal scenario

This scenario builds on the two scenarios described above. In this case, the defender has learned that the adversary has compromised the supply chain and has installed fake administrator accounts in network devices that are visible only when queried with a special modifier.

Assume that, as a result of the previous two scenarios, the defender has already collected information about the tactics, techniques and tools used by the adversary and has constructed a warning system that correlates certain Internet activity to specific power grid anomalies. The correlation system is believed to be accurate, particularly when dealing with this specific adversary.

Now assume that the adversary is upset about the outcome of the previous ensemble signaling scenario and decides to punish the defender by conducting additional cyber operations. The goal of the defender is deny benefits to the adversary and to impose a high cost on the adversary to deter it from conducting cyber operations. To achieve this goal, the defender establishes a reflexive signal, similar to launch on warning, that is triggered as soon as the defender's sensors detect an action by this particular adversary. The reflex is designed to corrupt the data stores on the adversary's operational networks, effectively crippling its capability to conduct cyber operations.

9. Conclusions

In 1996, Secretary of Defense William Perry outlined a strategy for managing armed conflict in the post-Cold War environment [23]. The first component of the strategy was to prevent threats from emerging. The second was to deter threats that emerged. The third, if prevention and deterrence

were to fail, was to defeat the threat using military force. Historically, signaling has been effective in implementing all three components involved in managing armed conflict. Clearly, signaling has an important role in managing conflict in cyberspace.

The signaling framework, which expresses adversary-defender interactions in terms of ensembles of adversary stimuli and defender signals, is both flexible and intuitive. It can model deterrence strategies in cyberspace as well as in other domains. Moreover, it provides an opportunity to formalize signaling plays to counter adversary actions based on defender goals. The scenarios involving cyber operations on the electric power grid illustrate the utility of the framework.

Note that the views expressed in this paper are those of the authors and do not reflect the official policy or position of the Department of Defense or the US Government.

REFERENCES

- [1] E. Poteat, The use and abuse of intelligence: An intelligence provider's perspective, *Diplomacy and Statecraft* 11 (2) (2000) 1-16.
- [2] S. Hildreth, *Cyberwarfare*, CRS Report for Congress, RL30735, Congressional Research Service, Washington, DC, 2001. www.au.af.mil/au/awc/awcgate/crs/rl30735.pdf.
- [3] United States Army, 2008 Army Posture Statement, Washington, DC, 2008. www.army.mil/aps/08/information_papers/transform/Cyber_Operations.html.
- [4] S. Gorman, Electricity grid in US penetrated by spies, *Wall Street Journal* (April 8) (2009).
- [5] S. Gorman, Electricity industry to scan grid for spies, *Wall Street Journal* (June 18) (2009).
- [6] W. Owens, K. Dam, H. Lin (Eds.), *Technology, Policy, Law and Ethics Regarding US Acquisition and Use of Cyberattack Capabilities*, National Academies Press, Washington, DC, 2009.
- [7] A. Berensen, As police use of tasers soars, questions over safety emerge, *New York Times* (July 18) (2004).
- [8] J. Langevin, M. McCaul, S. Charney, H. Raduege, (Co-Chairs); J. Lewis (Project Director), *Securing Cyberspace for the 44th Presidency*, Center for Strategic and International Studies, Washington, DC, 2008..
- [9] D. Patranabis, *Sensors and Transducers*, Prentice-Hall, New Delhi, India, 2004.
- [10] J. Richelson, *The US Intelligence Community*, Westview Press, Boulder, Colorado, 1999.
- [11] United States Strategic Command, *Deterrence Operations-Joint Operating Concept (version 2.0)*, Offutt Air Force Base, Nebraska, 2006. www.dtic.mil/futurejointwarfare/joc.htm.
- [12] R. Clarke, R. Knake, *Cyberwar: The Next Threat to National Security and What to do About it*, HarperCollins, New York, 2010.
- [13] T. Clancy, J. Gresham, *Submarine: A Guided Tour Inside a Nuclear Warship*, Berkley Books, New York, 2003.
- [14] R. Harper, A. Wiens, J. Matarazzo, *Nonverbal Communication: The State of the Art*, Wiley, New York, 1978.
- [15] P. Shenon, Can Obama shut down the Internet, *Yahoo News* (June 18) (2010).
- [16] Airforce-Technology.com, Israeli "e-tack" on Syria — Part 1, San Francisco, California, March 10, 2008. www.airforce-technology.com/features/feature1625.
- [17] L. Spitzner, *Honeypots — Tracking Hackers*, Pearson, Boston, Massachusetts, 2003.

-
- [18] J. Bradbury, S. Vehrencamp, *Principles of Animal Communication*, Sinauer Associates, Sunderland, Massachusetts, 1998.
- [19] National Research Council, *Letter Report from the Committee on Deterring Cyberattacks: Informing Strategies and Developing Options for US Policy*, National Academies Press, Washington, DC, 2010.
- [20] B. Blair, Alerting in crisis and conventional war, in: A. Carter, J. Steinbruner, C. Zraket (Eds.), *Managing Nuclear Operations*, Brookings Institution Press, Washington, DC, 1987, pp. 75–120.
- [21] H. Tuch, *Communicating with the World — US Public Diplomacy Overseas*, St. Martin Press, New York, 1990.
- [22] D. Sanger, Rice to discuss antiproliferation program, *New York Times* (May 31) (2005).
- [23] W. Perry, *Managing danger: prevent, deter, defeat*, *Defense Issues* 11 (13) (1996). www.defense.gov/Speeches/Speech.aspx?SpeechID=893.