The Digital Prisoner’s Dilemma: Challenges and Opportunities for Cooperation

Nadiya Kostyuk
Worldwide Cybersecurity Initiative, EastWest Institute
11 East 26th Street, 20th Floor
New York, NY 10010
nkostyuk@ewi.info; +1 212 824 4106

Abstract – Instead of neatly assigning ethics, morality, and responsibility to numerous cyber intrusions occurring all over world, this research analyzes the complexities of state relations in the realm of high-profile cyber attacks. Specifically, it utilizes an international relations prisoner’s dilemma to explain the complexities of cyber intrusions and how nation-states can deal with these new exigencies. This research looks at three possible scenarios between nation-states: 1) the powerful vs. the less-powerful; 2) the powerful vs. the powerful; and 3) U.S. vs. a “stepping-stone” nation. The conclusion reached in all three cases confirms that mutual cooperation between allies is the best option to deter or derail motivated offenders (state-sponsored or individual), especially now, when the formation of an international cybersecurity legal framework is still in a relatively inchoate stage.

Keywords— cyber attack; prisoner’s dilemma; international relations; powerful country; less-powerful country; and “stepping-stone” nation.

I. INTRODUCTION
Cyberspace has become a new, convenient battlefield, where states can operate anonymously due to a lack of international laws and regulations, the difficulty of attack origin attribution, the ability to use proxies to launch attacks, and states’ use of non-state actors. On this battlefield not only do national economies suffer significant damage (for instance, the U.S. lost $1 trillion worth of intellectual property as a result of corporate cyber espionage last year [1]), but every second individuals globally fall victim to various scams and viruses, such as the Nigerian 419 scam or the ILOVEYOU virus, due to the lack of both public awareness and law enforcement [2].

This research applies the classical international relations (IR) prisoner’s dilemma to cyberspace by examining the relations between two sets of countries (a powerful nation versus a less-powerful nation and a powerful nation versus another powerful nation) and provides plausible scenarios and responses in each case. Additionally, a third model—a stepping-stone country attacks the United States—is also evaluated for comparison. In the first two scenarios, the countries are divided into two categories based on the percentage of their GDP used for military expenditures. For the purposes of this research, the 15 countries that spend the highest percentage of their GDP on their military are identified as powerful countries and all other countries are less-powerful countries. Of the top 15, the 6 largest military spenders are the United States, China, Russia, the United Kingdom, Japan, and France; not surprisingly, five of them are veto-holders on the United Nations Security Council (UNSC) [3]. The stepping-stone model examines nations whose internet connectivity and lack of effective cyber policing capabilities provides motivated offenders with relative impunity. Even if a developed nation can track the attack origin, the stepping-stone nation has little to no means to track down cyber offenders within its borders. In addition to these models, observations and data from over sixty hours of semi-structured interviews with representatives from academia, various governments, and the press have been used to devise a universal approach for dealing with potential cyber tensions between various state actors. Since cybersecurity legislation is in its nascent stages in many countries and since most of them are developing their offensive and defensive cyber capabilities as well as domestic network infrastructure, the findings of this research are especially significant.

II. THEORETICAL FRAMEWORK: PRISONER’S DILEMMA
The prisoner’s dilemma is a concept used in game theory to explain why two individuals (or states) choose not to cooperate even though doing so would be beneficial to both. Merrill Flood and Melvin Dresher first formulated this concept in 1950 and in 1992, Albert Tucker described a situation in which two prisoners each have two options that determine their jail time. Table I illustrates Tucker’s example [4]. Table I demonstrates that each prisoner would receive a higher pay-off if he betrays the other. Lacking trust and fear of the other’s betrayal motivates both prisoners to testify against each other, even though the best option would be for them to cooperate. The realist theory of IR describes a similar situation between states that are often suspicious of

---

1 Definition of powerful and less-powerful country and a “stepping-stone” nation will be provided later on in the paper.
2 Stepping-stone countries—technologically unsophisticated states which often lack preventive cyber security measures in their police departments, laws that guard cyberspace, and the resources and institutional capacity to prevent attacks occurring in the online environment (Computer Crime Research Center 2010).
3 The researcher uses top military spenders instead of top purchasing power parity (PPP) as a measure of powerful countries, since the former have heavy internet saturation because they spend a significant amount of their GDP on building network infrastructure in addition to military spending. Top PPP countries often do not have comparable internet saturation rates among their population (e.g. Brazil and Saudi Arabia). The top PPP countries fit the hybrid model that will be explained later in the paper.
4 According to the Stockholm International research Peace Institute Yearbook 2013, the top fifteen nations that spend the most of their GDP percentage for military expenditures are: United Nations, China, Russia, the United Kingdom, Japan, France, Saudi Arabia, India, Germany, Italy, Brazil, South Korea, Australia, Canada, Turkey.
Estonia was chosen for this study because 1) Russia is the claim, many in Estonia believed the attacks to be of Russian origin, if not a deliberate attack by Moscow [12].

The case study of the Russian Federation and Estonia was chosen for this study because 1) Russia is the third in the world in military spending (after the U.S. and China), with a total of $90.7 billion U.S. [13]; and 2) Estonia was an ideal target for cyber attacks due to its status as an e-hub.

### TABLE I. TUCKER’S PRISONER’S DILEMMA

<table>
<thead>
<tr>
<th>Prisoner B stay silent (cooperates)</th>
<th>Prisoner B betrays (defects)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prisoner A stay silent (cooperates)</td>
<td>Each serves 1 year</td>
</tr>
<tr>
<td>Prisoner B serves 1 year</td>
<td></td>
</tr>
<tr>
<td>Prisoner A: 3 years</td>
<td></td>
</tr>
<tr>
<td>Prisoner B: goes free</td>
<td></td>
</tr>
<tr>
<td>Each serves 2 years</td>
<td></td>
</tr>
</tbody>
</table>

Waltz and Geico further explain that states react to the probability, not just the possibility, of threats posed by other states [7]. If a state simply reveals its capabilities, its opponents might start developing countermeasures; such is the case now as countries start developing their offensive and defensive cyber capabilities. Poland, for instance, intends to create a cyber unit within its army [8]. Moreover, some state actors have started using cyberspace to achieve their goals, as the internet provides them with plausible deniability via proxy servers and the use of non-state actors. This has proved to be alarmingly convenient, considering the conspicuous lack of laws and regulations at the international level aimed at punishing potential state perpetrators. Such anarchy in cyberspace and mistrust between countries create significant challenges in cooperation—best demonstrated by the IR’s prisoner’s dilemma.

### III. APPLICATION OF THE DIGITAL PRISONER’S DILEMMA

#### A. Powerful Nation v. Less-powerful Nation

On the ninth of May 2007, also known as the Soviet V-Day, Estonia experienced massive distributed denial-of-service (DDoS) attacks from computers all over the globe, including servers residing in the Russian Federation [9]. Several hours in, the attacks spread to a few major Estonian websites, causing most of them to be taken offline [10]. This digital malaise lasted for several weeks and the cyber assault against Estonia, by way of its sheer scope and scale, set a historical precedent: “never before had an entire country been targeted on almost every digital front all at once, and never before had a government itself fought back” [11]. Though there was no substantial evidence to validate the claim, many in Estonia believed the attacks to be of Russian origin, if not a deliberate attack by Moscow [12].

The case study of the Russian Federation and Estonia was chosen for this study because 1) Russia is the third in the world in military spending (after the U.S. and

Based on the amount of avoided questions and uncomfortable silences by some of the interviewed Estonian experts, it seems that Estonia was stuck in the classic prisoner’s dilemma—a situation in which a country can be a victim, but cannot afford to identify its victimizer. Despite the general consensus of Estonian and Russian interviewees on the difficulty of proving Russia’s responsibility for the Estonian cyberattacks, such a unified view creates an area for potential and rather challenging research on whether the interviewees, who represent a wide spectrum of contemporary cyber expertise, genuinely did not know who was behind the attacks or simply did not want to make accusations or admit guilt or failure. If the Estonian attacks did in fact originate in Russia and the Russian state is not responsible, then it could be viewed as either a failure of the Russian state’s ability to police its cyberspace or a deliberate use, by Moscow, of the hacking community, specifically the youth group NASHI, for the state’s interest. This situation is representative of every country’s struggle with the prisoner’s dilemma in the high stakes, high drama game of international relations.

When in the prisoner’s dilemma, the less-powerful country has two options: it can either refuse to admit that a powerful nation can effortlessly hack into its infrastructure without fear of retribution by the less-powerful state, or it can purposely express its weakness to get help from powerful allies such as the EU, NATO, or individual member states therein. Moreover, a less-powerful country will generally be unable to prove that a powerful nation is the perpetrator due a lack of funding for origin attribution in their cyberspace, politico-economic considerations, or a simple lack of courage to accuse the powerful nation outright. Making accusations sans concrete evidence can result in a less-powerful nation to provoke a powerful nation with far more financial and military resources into willfully and brazenly violating the smaller states’ sovereignty as a show-of-force. The worst-case scenario for the victim-country would be to wrongfully accuse a powerful nation of sovereignty violations and, thus, inadvertently provoke a potentially severe diplomatic crisis [14].

5 These experts spoke to me on the condition of anonymity.
Despite being the best option, cooperation is very unlikely between powerful and less-powerful nations. One possible explanation for why a powerful nation is not likely to cooperate with a less-powerful nation is that the former’s political leaders might appear weak to its citizens and on the international stage [15]. Geico emphasizes the point that nation-states decide who will benefit more from this cooperation before they choose whether or not to cooperate [16]. Thus, a powerful country will usually lack incentive to cooperate with a less-powerful country as the latter has little to offer for this cooperation. For instance, in the case of the Estonian cyber attacks, the Russian Federation was not interested in cooperating because it had nothing to do with its interests [17]. Irina Laguna, a senior broadcaster at Radio Free Europe/Radio Liberty, an international news agency funded by the United States, considers visokomerije (arrogance) as the main reason for Russia’s refusal to help Estonia investigate the 2007 cyber attacks [18], despite the fact that the two countries have a Mutual Legal Assistance Treaty [19]. Not only does such a refusal make a proper investigation of the attacks impossible, while creating a sense that Russia might be behind the attacks [20], but it also violates international law. On the other hand, despite all circumstantial evidence Estonia did not officially accuse Russia (rather it shared its suspicion with the U.S.) [21] and chose to seek assistance from NATO in developing stronger cybersecurity protection measures [22]. Less-powerful nations will most likely follow Estonia’s example in cooperating with friendly and trustworthy powerful unions that have far more resources to develop cyber capabilities, implement cyber protective measures, and prevent future attacks in the online environment. The idea of less-powerful nations cooperating with each other is quite appealing; however states that still lack cyber capabilities will most likely seek protection from states that have these important resources. On the other hand, considering how fast cyber offensive and defensive capabilities are being developed, one can expect that less-powerful nations will start cooperating with each other quite soon as keeping up with the tech race will likely be too costly countries with limited financial resources.

B. Powerful Nation v. Powerful Nation

The U.S. and China were chosen for this case study to represent two powerful nations because 1) they are the top two countries of military expenditures [23] and 2) they often accuse each other of committing cyber attacks.

### TABLE III. PRISONER’S DILEMMA FOR THE U.S. AND CHINA

<table>
<thead>
<tr>
<th>Countries increases</th>
<th>Does not cooperate</th>
<th>Unlikely Scenario:</th>
<th>Highly Likely Scenario:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1) The attacks escalate; 2) U.S. relies on its adept domestic and international law enforcement arms; 3) China appears incapable of policing its cyberspace, making it vulnerable to internal attacks and eventually is forced to cooperate with the States.</td>
<td>1) The attacks escalate; 2) U.S. relies on its adept domestic and international law enforcement arms; 3) China appears incapable of policing its cyberspace, making it vulnerable to internal attacks and eventually is forced to cooperate with the States.</td>
</tr>
<tr>
<td>China (willingness to talk)</td>
<td>Co-operates</td>
<td>1) Individual hackers are punished; 2) Trade between the two nations continues</td>
<td>1) U.S. continues experiencing losses in its intellectual property; 2) The U.S. could try applying sanctions against China; 3) Mutual Legal Assistance Treaty is worthless?</td>
</tr>
<tr>
<td></td>
<td>Does not cooperate</td>
<td>Highly Unlikely Scenario:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1) U.S. denies responsibility; 2) U.S.-China relations worsen; 3) Trade declines causing severe economic losses in the States; 4) The number of cyber attacks coming from both</td>
<td></td>
</tr>
</tbody>
</table>

6 Specifically the U.S. spends $682 billion U.S. and China spends $249 billion U.S.  

In a short timeframe, China has developed a long history of cyber espionage, becoming active in many parts of the world. China has established two network spy stations in Cuba to monitor U.S. internet traffic and Department of Defense communications [24], conducted Titan Rain, an incident involving an “extraction of between 10 to 20 terabytes of data from the Pentagon’s unclassified network” [25], hacked into the computers of German Chancellor Angela Merkel and former U.S. Secretary of Defense Robert Gates, and conducted GhostNet, which infiltrated approximately 1,300 computers at various embassies around the world. China continues hacking U.S. companies either to steal their intellectual property or because they find it hard to believe that American companies are transparent and honest in doing business with them. The Mandiant report, released in March 2013, highlights the peculiarities of over a decade of Chinese hacks into U.S. infrastructure, government, ministries, and the financial sector with the main purpose of stealing information [26]. As a response to this report, China blamed the U.S. for continuously hacking Chinese computers [27]. As discovered through Edward Snowden’s revelations [28], the United States assumes the same lack of business integrity as China and, therefore, also has incentive to hack the country.

Insecurity and fear of being attacked are the key motivators behind these attacks. Kenneth Geers, the first U.S. Representative to the Cooperative Cyber Defense Center of Excellence (CCD CoE) in Tallinn and a cybersecurity expert at FireEye, provides a historical precedent as an explanation to this conundrum by stating that governments were, are, and will be spying on each other, and the only difference is that now they are able to spy using a computer [29]. Countries’ dependency on each other makes a response rather complex. Geers provided an example of this dependency: the U.S. and Australia are dependent on their Chinese relations and therefore do not take radical action against perceived Chinese malfeasance [30]. Even though the U.S. was able to find an exact location of hackers in China, it still cannot blame the Chinese government openly or outright. Moreover, unlike the U.S. (despite the existence of PRISM), China has complete control over its internet infrastructure (especially due to government-mandated hardware censorship technology), leading to the conclusion that China either has control or is unaware of the hackers in its domain. Rather than air this dirty laundry in the light of day, Americans continue to ensure the Chinese through political and economic cooperation, tying the fate of China to that of the U.S., a textbook prisoner’s dilemma if there ever was one.

Despite the mutual lack of trust, especially in this newly developing field, both the U.S. and China have realized

the importance of cooperation. Escalating cyber attacks may have devastating economic collateral damage that leads to trade decline between the two countries with the United States taking the hardest hit and continuing cyber attacks might lead to a cyber war—a polarizing hypothesis that has been both vehemently disputed and strongly purported. Therefore, the U.S. has utilized informal talks to address the hacking matter behind closed doors, away from the microphones and cameras of the fourth estate. Specifically, China and the U.S. are making some progress on cybersecurity cooperation. For instance, the U.S.-China Strategic and Economic Dialogue is a step in the right direction because it increases mutual trust and decreases suspicion between the two countries [31].

Such a situation is applicable to any two powerful nations that are forced to cooperate with each other in developing norms and regulations. A lack of cooperation could cause devastating results to both nations, including problems with infrastructure and collateral damage. The EastWest Institute has established a significant foundation for a common framework by defining cyber terminology between Russia and the U.S. [32] and is planning to start a similar round between the U.S. and China. These discussions will serve as a strong foundation for legal agreements that will be developed in the near future to mitigate any misbehavior in cyberspace. This excellent example should be followed by other nation-states.

C. “Stepping-Stone Nation” (SSN) Model

It would be interesting to examine the hypothetical scenario in which cyber offenders launch attacks against the U.S. from places that serve as their safe havens and provide them with plausible deniability. Even if a developed nation can track the attack origin, the stepping-stone nation has little to no means to track down cyber offenders within its borders. Moreover, these nations rarely have extradition treaties with other nations, as demonstrated by the ILOVEYOU virus in the Philippines [33].

TABLE IV. PRISONER’S DILEMMA FOR THE U.S. AND A “STEPPING-STONE” NATION

<table>
<thead>
<tr>
<th>Stepping-stone Nation (SSN)</th>
<th>Cooperates</th>
<th>Does not cooperate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unlikely Scenario: 1) Individual hackers are punished</td>
<td>Highly Unlikely Scenario: 1) SSN is willing to talk; 2) U.S.-SSN relations worsen; 3) The number of cyber attacks coming from both countries increases</td>
<td></td>
</tr>
<tr>
<td>Does not cooperate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likely Scenario: 1) The U.S. continues experiencing cyber attacks;</td>
<td>Highly Likely Scenario: 1) The attacks escalate from both parties; 2) The U.S. could try applying sanctions against a SSN.</td>
<td></td>
</tr>
</tbody>
</table>

1. Though there is obviously durable economic relationship between the U.S. and China, there is still much insecurity between these two countries and as the December 2013 airspace crisis [http://www.nytimes.com/2013/08/05/world/asia/december-2013-airspace-crisis-time-dispute.html] showed the economic relations might not be the underlying condition for cooperation.

2. World-renowned expert on security system design Marcus Ranum

3. Including former U.S. National Coordinator for Security, Infrastructure Protection, and Counter-terrorism at the State Department Richard Clarke

Table IV demonstrates a scenario in which an SSN and a powerful country, such as the U.S., attack each other. Elements from the first and second scenarios are combined as the powerful country will either be forced to cooperate with the SSN due to an escalating nature of the attacks or will use international community assistance to apply sanctions against this nation. This scenario can be developed further to include less-powerful states that have the potential to become technologically superior in a few decades. Specifically, the Nigerian 419 scam has demonstrated that motivated offenders do not have to come from wealthy, technologically-sophisticated nations to do great harm globally.

IV. DISCUSSION: SOLVING THE DIGITAL PRISONER’S DILEMMA

The best solution to the digital prisoner’s dilemma is cooperation. Due to the lack of trust and transparency, different views on information security exist (United States versus China versus Russia) and are utterly disparate and incompatible with each other, resulting in massive gaps in the ability to adjudicate cybercrime (state-sponsored or individual). In her book Anatomy of Mistrust, political scientist Deborah Larson argues that mutual mistrust may create failures in cooperation, resulting in expanding the already sizeable gap “which works to perpetuate [the] cycle by worsening the dynamic that already existed at the outset” [34]. Therefore, the best recommendation is cooperation as nations have more to gain by working together than by fighting as individuals.

Government cooperation in cybersecurity is currently in its nascent development and as of now, only non-governmental organizations (NGOs) and think tanks have taken the initiative to build bridges between various countries. The EastWest Institute, for instance, holds annual cybersecurity summits using Track 1.5 or Track 2 dialogue during which countries can learn about each other in order to foster cooperation. Though fruit has yet to be born of these dialogues, it is hopeful that countries will soon be able to reach a consensus on cybersecurity matters as they realize that there is more to be gained by working together than by using the alternatives previously mentioned. For instance, the U.S. and Russia have been working on creating common cyber terminology [36], and the U.S. and China have informally agreed to develop mutual rules of engagement in terms of cyber strikes [37]. This approach has a promising future in the realms of cybersecurity and international relations since avoiding conflict is generally the preferred outcome regardless of the IR model one has an affinity for.

Most small nations have demonstrated a lack of fiscal resources to go toe-to-toe with an advanced persistent threat (APT) launched from the U.S., Russia, or China. Small nations should look forward to cooperating with each
other in order to bolster themselves collectively in the face of greater threats, state sponsored or individually enacted, and then work with larger nations as a collective. Though these case-study nations are already affiliated with groups (notably NATO and the EU), the lack of a collective cyber-defense paradigm and the continuing zeal for domestic internet sovereignty makes these collective organizations currently ineffective in mitigating APTs from public or private actors. This could change with some sacrifices by affiliated nations and with a renewed drive for collective defense against twenty-first century threats.

V. TAKEAWAYS AND AVENUES FOR FUTURE RESEARCH

This research’s findings are significant for the field of international relations as it applies the classical prisoner’s dilemma to cyberspace, a new dimension with many exigencies and a few solutions. The common lesson derived from the described scenarios—international cooperation is important for a safer cyber world—is useful for other nations, developed and developing, rich and poor, because the World Wide Web erases any physical borders between nations and levels the playing field between powerful and less-powerful states. In light of recent cyberattacks across the globe, internet-sophisticated nations proved to be as vulnerable to attacks as nations that are in the inchoate phase of their internet development.

In addition to significant technical expertise, the exigencies of cybersecurity in the twenty-first century now involves and necessitates more policy than in the past. This research can be expanded to include accurate data from nation-states to as victimization rates, known domestic APTs, and public discussion of the merits of non-classified measures being taken to protect national cyberspace. Additionally, data from these nations can be analyzed through a statistical game theory model in which statisticians and political scientists can draw empirical predictions. On the technical side, computer engineers can use the policy environment described here in order to advance internet protocols for identifying and preventing proxy attacks. Interviews with computer specialists and engineers will add value to this research by giving a well-rounded strategy for the future of cybersecurity as will peer-reviewed empirical research on hardware- and software-based approaches to cybersecurity.

VI. CONCLUSION

“If you are defending cyberspace, you're already too late. If you do not dominate cyberspace, you cannot dominate it in other domains. If you are a developed country [and you are attacked in cyberspace], your life comes to a screeching halt”[39].

Over the last two decades, the internet has created new opportunities, vulnerabilities, and a new battlefield for future conflict. Possessing resources and capabilities, powerful countries have started using this new avenue to gain an advantage over each other and maintain their advantage over less-powerful nations. The scenarios described in this paper demonstrate a classic prisoner’s dilemma. Less-powerful countries do not have the ability to accuse their victimizer; instead they either have to cooperate with the offender or seek support from powerful allies. Powerful nations, on the other hand, continue cyber attacks against each other, while making small steps in private towards cooperation. Despite the diversity in cultures, traditions, and opinions, states worldwide should seek to cooperate with each other in addressing the complexities of cyberattacks and establishing the means to mitigate them. Therefore, worldwide cyber cooperation initiatives should be undertaken to mitigate future cyber attacks because until this is achieved, cyber offenders will flourish and attacks will increase in both number and severity and weak nations will continue to be cyber-safe havens for motivated offenders. If the global community does not cooperate, then the forward progress of globalization and its promise of peaceful development will remain subject to drastic, and eventually catastrophic scrreeching halts.

REFERENCES

Geico, supra n 5.


Lagunina, Irina, Personal interview, 24 June 2013


Wikileaks source, a US cable date 6/4/07.

Michaels, supra n 17.

"Military Expenditure Database," supra n 3.


Ibid.


Ibid.


Department of the Army, supra n 2.


Rauscher and Yaschenko, supra n 32.


Montville, supra n 35.

The Director of the Air Force Cyberspace Operations Task Force, as quoted by Clarke & Knake, 38, supra n 24.