Journal of the Association for Information Systems (2018) **19**(3), 152-181
doi: 10.17705/1jais.00488

RESEARCH PAPER

ISSN 1536-9323

Preventing State-Led Cyberattacks Using the Bright Internet and Internet Peace Principles

Young Yung Shin[1], Jae Kyu Lee[2], Myungchul Kim[3]
[1]Korea Advanced Institute of Science and Technology, inewhero@kaist.ac.kr
[2] Xi'an Jiaotong University, Carnegie Mellon University, and KAIST, jklee@kaist.ac.kr
[3]Korea Advanced Institute of Science and Technology, mck@kaist.ac.kr

Abstract

The Internet has engendered serious cybersecurity problems due to its anonymity, transnationality, and technical shortcomings. This paper addresses state-led cyberattacks (SLCAs) as a particular source of threats. Recently, the concept of the Bright Internet was proposed as a means of shifting the cybersecurity paradigm from self-defensive protection to the preventive identification of malevolent origins through adopting five cohesive principles. To design a preventive solution against SLCAs, we distinguish the nature of SLCAs from that of private-led cyberattacks (PLCAs). We then analyze what can and cannot be prevented according to the principles of the Bright Internet. For this research, we collected seven typical SLCA cases and selected three illustrative PLCA cases with eleven factors. Our analysis demonstrated that Bright Internet principles alone are insufficient for preventing threats from the cyberterror of noncompliant countries. Thus, we propose a complementary measure referred to here as the Internet Peace Principles, which define that the Internet should be used only for peaceful purposes in accordance with international laws and norms. We derive these principles using an approach that combines the extension of physical conventions to cyberspace, the expansion of international cybersecurity conventions to global member countries, and analogical international norms. Based on this framework, we adopt the Charter of the United Nations, the Responsibility of States for Internationally Wrongful Acts, Recommendations by the United Nations Group of Governmental Experts, the Tallinn Manual, and Treaty of the Non-Proliferation of Nuclear Weapons, and others as reference norms that we use to derive the consistent international order embodied by the Internet Peace Principles.

Keywords: Bright Internet, Cyberattack, State-Led Cyberattack, Preventive Cybersecurity, Internet Peace Principles

1 Introduction

Information Cyberattacks present serious threats to national infrastructure and defense systems as well as to the private sector. Cyberattacks can be defined as "actions taken through the use of computer networks to disrupt, deny, degrade, or destroy information resident in computers and computer networks, or the computers and networks themselves" (Joint Chiefs of Staff, 2009, p. 111). Such threats not only have private-led origins, but also state-led origins.

In this paper, we propose a framework that can prevent state-led cyberattacks (SLCAs), which state agencies themselves and/or employed individuals or companies conceive of and operationalize. SLCAs are capable of attacking not only military and governmental systems but also critical civilian infrastructure, such as financial systems, telecommunication systems, energy systems, transportation systems, and private corporations. In

contrast, private-led cyberattacks (PLCAs) may also threaten governmental systems, as well as private companies, homes, individuals, and civilian communities, but do so mainly to gain illegal economic advantages through ransomware, phishing and/or stolen credit cards. By private actors we mean nonstate actors, including individuals, ordinary citizens, script kiddies, hackers, "hacktivists" (such as Anonymous), "patriot hackers" (a student-run cybersecurity group formerly known as the Electric and Computer Hacking Organization), cyberinsiders, cyberterrorists, malware authors, cyberscammers, organized cybercriminals, and corporations (Sigholm, 2013).

To characterize the nature of SLCAs, we collected seven typical SLCA cases and contrasted them with three illustrative PLCA cases. To compare and contrast their differences and commonalities, we organized the cases according to the following eleven perspectives: attack purposes, targets of attacks, origin country, attack means and methods, attack routes, timing, duration, preparation period, investigation period, consequences, and applicable laws. We found that the specific actors, purposes, targets, attack timing, means, methods, and circumstantial evidence pertaining to SLCAs differ significantly from those of PLCAs despite the fact that the basic technologies for both are nearly identical. Because attacking countries do not admit their responsibility, origin traceability technology alone cannot identify the origins of SLCAs. Thus, an analysis of other evidence, such as repeated introductions of malware from a particular country and/or circumstantial evidence, should be considered as well.

Assuming that the sources of anonymous cyberattacks from global origins are uncontrollable, current cybersecurity systems primarily attempt to defend their own systems reactively. To overcome such a limitation, the concept of the Bright Internet was recently proposed as a framework for preventive security that makes the origins of malicious cyberattacks transparent, traceable, and identifiable. In the current paper, we adopt the Bright Internet framework for preventing SLCAs by adopting its five principles: origin responsibility, deliverer responsibility, identifiable anonymity, global collaboration, and privacy protection (Lee et al., 2018, Lee, 2015).

By adopting the Bright Internet framework, the premise is that transparency can deter the generation of SLCAs from member countries who conform to the principles of the Bright Internet. Among member countries, participating governments will monitor for the malicious emission of cyberattacks within their own countries according to established responsibility chains which can be inherited from the international law on the Responsibility of States (2001). However, the origins of SLCAs may not be fully identifiable if cyberterror countries do not honestly report the malicious origins, especially when plotting SLCAs.

Moreover, not all states may agree to become members of the Bright Internet. Therefore, it is necessary to add a complementary measure, which we term the **Internet Peace Principles**, which prohibits the use of the Internet as a weapon for attacking other countries or as a means of detoured malicious attacks. By combining the five principles of the Bright Internet with the Internet Peace Principles, we aim to design a framework of preventive security that can deter SLCAs.

To derive the concept and practice of the Internet Peace Principles, we adopt a combination of the following three approaches: the extension of physical conventions to cyberspace, the expansion of international cybersecurity conventions to global member countries, and analogical international norms. We adopt the Charter of the United Nations (UN Charter), draft articles on the Responsibility of States for Internationally Wrongful Acts (Responsibility of States), the Geneva Convention IV (Geneva Convention), and the Hague Convention IV (Hague Convention) in order to extend the traditional convention to cyberspace. We adopt the UN Group of Governmental Experts (UN GGE), the Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations[1] (the Tallinn Manual) through an invitation by the NATO Cooperative Cyber Defence Center of Excellence, and the Council of Europe Convention of Cybercrime (CECC) for the expansion of international cybersecurity conventions to global member countries. In addition, for the derivation of similar principles from analogous international norms, we adopt the Outer Space Treaty and the Treaty on the Non-Proliferation of Nuclear Weapons (NPT). From these existing conventions and the Bright Internet principles, we derive the Internet Peace Principles which can contribute to the prevention of SLCAs.

The remaining sections of this study are organized as follows. In Section 2, we review two categories of literature related to this paper: security and privacy research in information systems literature and referential international conventions. In Section 3, we review the principles of the Bright Internet to explain how they can deter cyberattacks through the transparent traceability and identifiability of attacking origins while protecting the freedom of expression and privacy of innocent netizens. We review five principles

[1] The Tallinn Manual 2.0, published in 2017 contains the Tallinn Manual 1.0 that did not address legal problems of cybersecurity outside armed conflicts published in Fleck (2013).

of the Bright Internet in relation to the characteristics of SLCAs. In Section 4, we analyze seven SLCA and three PLCA cases along eleven perspectives to determine the differences and commonalities between them. Section 5 reviews the benefits of the Bright Internet principles for Bright Internet Global Organization (BIGO) member countries. However, we also recognize their limitations if cyberterror countries are not signatories to the BIGO Agreement. In order to develop a complementary measure, we use current conventions to derive the Internet Peace Principles in Section 6. Section 7 concludes with a discussion of the future research agenda.

2 Literature Review

We review two categories of literature.

2.1 Security Research in the Information Systems Literature

Lee et al. (2018) reviewed the literature on information systems security research with two-dimensional perspectives toward both the target entity (e.g., individual, organization, society) and research methodology (e.g., behavioral science and design science). According to this review, most security research in information systems journals has dealt with behavioral studies of individuals and organizations.

Anderson and Agarwal (2010), Wang et al. (2015), Chen and Zahedi (2016), Johnston and Warkentin (2010), and Steinbart et al. (2016) conducted individual behavioral research on the topics of behavioral security, protection motivation, and information security threats. Herath and Rao (2009a, 2009b), Bulgurcu et al. (2010), D'Arcy et al. (2009), Willison and Warkentin (2013), Mitra and Ransbotham (2015) and Hsu et al. (2015) conducted organizational behavioral research on the topics of intraorganizational information security policy/technology and human-related issues.

Oetzel and Spiekermann (2014) reviewed the multidisciplinary nature of privacy research, and Belanger and Crossler (2011) conclude that most privacy research in the information systems domain concentrates on individual-level issues. Therefore, Belanger and Crossler (2011) and Pavlou (2011) stress the need for research on design and action at the societal level.

In contrast to the existing literature in information systems journals, research on preventive security against SLCAs is unusual, and this paper is likely the first attempt. From the perspective of behavioral research, the Bright Internet needs experimental studies at the societal level concerning netizens' behavior in order to justify social norms pertaining to the Bright Internet and the Internet Peace Principles. The research subjects could be individuals, organizations, and/or countries; security behavioral research would certainly be expanded by this type of research. Various empirical studies could emerge following the actual deployment of the Bright Internet system and ensuing test bed.

However, during the early stages of Bright Internet research, more attention should be paid to the design science aspect of the global information system infrastructure in terms of technologies, policies, and international collaborations. Lee et al. (2018) proposed the design science perspective of Bright Internet development, and the frameworks of Gill and Hevner (2013), Hevner et al. (2004), and March and Smith (1995) are useful for validation purposes.

2.2 Referential International Conventions

There exists no established global convention for peace in cyberspace similar to our Internet Peace Principles. The extant conventions were established long ago for a physical-space context and have not been effectively applied to cyberspace. The scope of current research on the cyberconvention is fragmented and often only regionally applicable. As such, a comprehensive set of principles is needed that is consistent with current social norms and that considers the characteristics of the Internet. Whereas Lee et al. (2018) adopted prevention motivation theory and analogical social norms to justify five Bright Internet principles, the Internet Peace Principles adopt three approaches, as discussed above. We review the current conventions below.

For the extension of the traditional conventions on cyberspace, we reviewed the UN Charter (1945), the Responsibility of States (2001), the Geneva Convention IV (1949), and the Hague Convention IV (1907).

The **UN Charter** article 1 in chapter I ("Purposes and Principles") states that the purpose of the United Nations is "to maintain international peace and security," which is also a key foundational element of the Internet Peace Principles. Article 1 also states the aim "to take effective collective measures for the prevention and removal of threat to the peace and for the suppression of acts of aggression or other breaches of the peace." Article 51 in chapter VII ("Action with Respect to Threats to the Peace, Breaches of the Peace, and Acts of Aggression") refers to the practice of self-defense and countermeasures: "UN Security Council may take . . . demonstrations, blockades, and other operations by air, sea, or land forces of Members of the United Nations" according to article 42 in chapter VII. These provisions should be extended to the Internet peace and cyberattack contexts.

The **Responsibility of States** defines that "every internationally wrongful act of a State entails the international responsibility of that State" (article 1). It also specifies that wrongful acts include wrongful actions and omissions of duty (article 2), and defines the person or entities of government authority (article 5) which can be extended to include software that is programmed by governmental persons. Articles 21 and 22 specify the conditions to justify self-defense and countermeasures, and articles 34-37 define the types of reparations of injuries and damage as restitution, compensation and satisfaction. These articles could easily be extended to the Internet context.

The **Geneva Convention IV** preamble and articles prohibit attacking civilian and civilian objects, and the **Hague Convention IV** regulates the unlimited use of means and methods of warfare, both of which could be extended to the cyberattack context.

For the expansion of international cybersecurity conventions, representative cases selected here are the United Nations Group of Governmental Experts (UN GGE), the Tallinn Manual, and the Council of Europe Convention on Cybercrime (CECC; 2001). Their initiatives must be expanded to global members to make them effective among all Internet user countries.

The **UN GGE** consists of twenty countries and is based on geographical distribution; it includes cyberpower countries such as the USA, China, Russia, France, the UK, and Germany. The member countries of this convention have met since 2004 and created reports recommending norms and principles. According to the 2015 UN GGE report, the recommendation includes the following limiting norms: States "should not knowingly allow their territory to be used for internationally wrongful act using ICTs" and "should not conduct or knowingly support ICT activity that intentionally damages or otherwise impairs the use and operation of critical infrastructure" (United Nations, 2015, p. 2). This refers to the prohibition of SLCA acts. The recommendation also prevents the proliferation of malicious ICT; this supports the Bright Internet principles of origin responsibility and deliverer responsibility. The UN GGE also recommends the positive duties of cooperation by exchanging information and prosecuting terrorists and the criminal use of ICT (Park & Chung, 2016), which can be seen as the basis of the global collaboration principle for the Bright Internet.

The **Tallinn Manual** is the only document that specifies the international laws applicable to cyberwarfare and cyberoperations. The manual provides rules about state actors' responsibilities as well as their jurisdictions and countermeasures during cyberconflicts. The manual was produced by the independent International Group of Experts as a nonbinding document. The NATO Cooperative Cyber Defence Center of Excellence produced the Tallinn Manual 1.0 in 2013, and updated it in 2017 as the Tallinn Manual 2.0 which identified the international laws applicable to cyberwarfare and cyberoperations in 154 rules (Schmitt, 2017).

The manual (2017) stipulates various regulations reflecting the characteristics of cyberspace, including diplomatic and consular law (rules 45-54), international telecommunication law (rules 61-64), the law of the sea (rules 45-54), the air law (rules 55-57), and the space law (rules 58-60). It specifies sovereignty (rules 1-5), and states' due diligence (rules 6-7), and jurisdiction (rules 8-13). In this regard, the manual mentions states' responsibility and internationally wrongful acts. Rules 20-26 prescribe the cyberattacked state's entitlements to seek countermeasures. In particular, it details the offending state's responsibility concerning assurance, guarantees, and reparation (rules 27-29). The manual elaborates cyberoperations not regulated by international law, such as peacetime cyberespionage (rules 32-33). In addition, it extends the current international regulations to cyberspace, including the principle of nonintervention (rules 66-67) and peaceful resolution of conflicts (rule 65). The manual describes the prohibition of the use of force as a critical issue (rules 68-70) and explains when and how this principle is applied. Moreover, in the same context, it specifies when cyberactivities constitute an armed attack (rule 71) and hostility (rules 86-130) and therefore how and when cyberattacked states should conduct self-defense (rules 71-75) which is a cyberextension of the Geneva Convention. It also includes the prohibition of attacking civilians and their facilities (rules 94-102) and the use of cyber means and methods that cause unnecessary injury or suffering (rules 103-105) during hostilities. It prohibits attacking places of worship and nuclear electric power generating stations (rule 109). It also notes the use of ICT for the enforcement of naval and aerial blockades, and discusses the validity of cyberblockades (rule 128). However, it explicitly prohibits collective punishment by cybermeans (rule 144), unlike the UN Charter which seeks collective measures for the prevention and removal of threats. These stipulations need further study because collective cyberblockades may occasionally be necessary to prevent cyberattacks from cyberterror countries.

The **CECC** is an international agreement founded in 2001 for the purpose of addressing Internet and computer-network crime through establishing common criminal policies, procedures, and methods among member countries. The CECC refers to the need to pursue a common criminal policy to protect

society against cybercrime. It recognizes "the need for cooperation between states and private industries in combating cybercrime" (CECC, 2001, p. 2) and holds that effective measures against cybercrime require strong, timely and well-functioning international cooperation in criminal matters. However, the CECC does not cover issues associated with SLCAs.

To derive analogical international norms, we selected two relevant treaties: the Outer Space Treaty and the NPT under the auspices of the International Atomic Energy Agency (IAEA). The Outer Space Treaty was adopted in 1967, with the signers agreeing that the benefits of space exploration should be bestowed on all of humankind (Porras, 2006). The treaty recognizes that outer space shall be the province of all humankind—cyberspace is also virtually borderless and should thus adopt the same spirit.

In the NPT (1968) and the vision of U.S. President Dwight D. Eisenhower's "Atoms for Peace," we find support for prohibiting the proliferation of fundamental threats to cybersecurity, such as the cases demonstrated in Section 4, through adopting the Bright Internet and Internet Peace Principles. Likewise, the proliferation of the malicious production and distribution of cybermalware should also be prohibited. It is noteworthy that the IAEA was established as a body researching nuclear safety, security, and verification as well as the development of peaceful uses of nuclear power. Similarly, we need a global agent, like the BIGO, that is devoted to research on and the deployment of the Bright Internet in cooperation with existing international organizations, as described in Section 6.

3 The Principles of the Bright Internet

The principles of the Bright Internet proposed by Lee in 2015 were more recently updated by Lee et al. (2018). This section reviews how the principles of the Bright Internet are designed to identify the origins of cyberattacks while ensuring the freedom of expression and protection of privacy for innocent netizens. To attain these goals, we adopt five principles, depicted in Figure 1 below: the principles of origin responsibility, deliverer responsibility, identifiable anonymity, global collaboration, and privacy protection. Lee et al. (2018) argued for the need for these principles based on the perspectives of prevention motivation and analogical social norms. Below, we review the definitions of these principles and discuss their applicability to SLCAs.

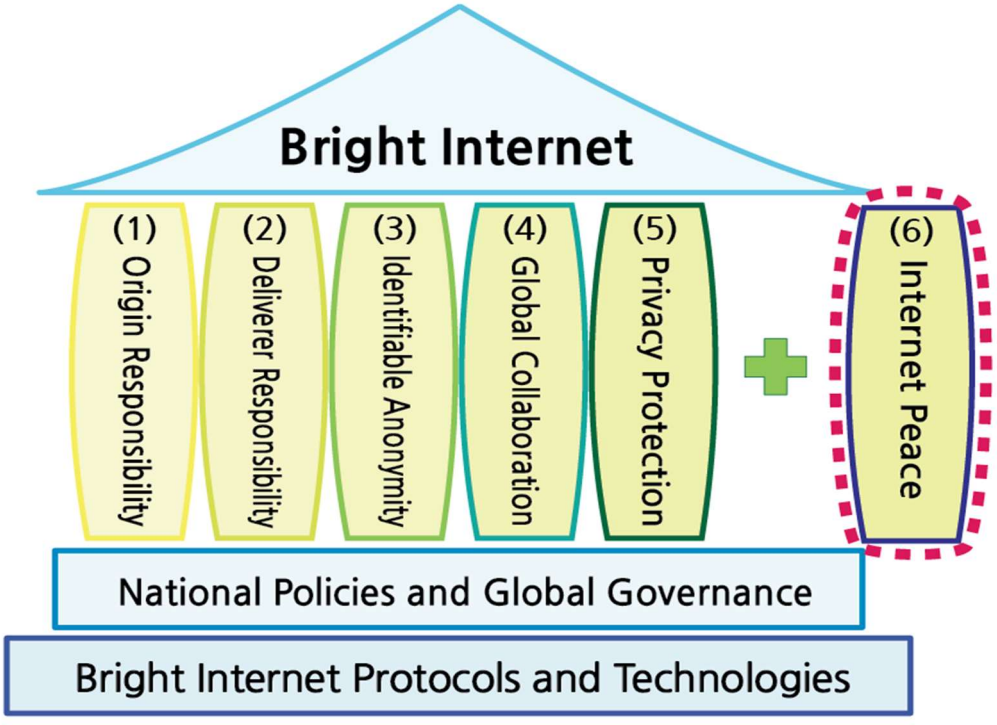

Figure 1. Bright Internet and Internet Peace Principles

The **principle of origin responsibility** states that the "offensive originators of malicious codes and illegal hacking should be responsible for the consequence of their malicious behaviors" (Lee et al., 2018, p. 71). This is basically identical to the principle of state responsibility. To realize this principle, computer users should not only be concerned about protecting their own systems but should also take care not to attack others from their computers. Layers of origin responsibility refer to the individual users, servers, companies, and countries associated with the cyberattacks. However, with regard to illegitimate SLCAs, the responsible country may hide the individuals and servers who commit attacks; hence, responsibility for this type of unidentified origin should be borne by the country in which it emerges. In addition, the attacking government may utilize computing resources outside of their own country, as demonstrated in Section 4. This is why the principle of deliverer responsibility is essential to offset the loophole of origin responsibility.

The **principle of deliverer responsibility** means that "compromised computers or Internet service providers who are involved in the delivery process of cyberattacks, even unintentionally, should cooperate to prevent the delivery of identifiable harms to users at the destination" (Lee et al., 2018, p. 72). To realize this principle, computer users should avoid attacking others as zombie computers. For instance, when a malicious DDoS attack is made on U.S. domain name servers, U.S. citizens and internet service providers should seek to avoid acting as zombies for the enemy. Through the effort of participative prevention, global netizens can protect each other and maintain peace in their countries.

If attackers abuse the resources of third countries to stage detoured cyberattacks, the citizens of such a country should take preventative measures to avoid being compromised and unwittingly participating in attacks on friendly countries. This should be the ethical and legal standard for global safety of the Internet around the world. Both PLCAs and SLCAs should be prevented by imposing this principle on those countries who agree with it. A country that does not agree with this type of mutual protection should be regarded as a potential cyberterror country and treated differently. This is why a collective measure is necessary.

Even when it is possible to identify the origin IP address, malicious attackers will use pseudonyms to hide their actual names. This is why the **principle of identifiable anonymity** is necessary, which means that "The real name or equivalent identity of the criminal origin should be identifiable in nearly real time in the context of a valid search warrant, while the voluntary anonymity of innocent netizens should be preserved" (Lee et al., 2018, p. 73). There are a number of approaches that can be used to implement the principle of identifiable anonymity, but cyberterror countries that commit SLCAs will not be interested in abiding by this principle. This is why the Internet Peace Principles need a sanction mechanism, as described in Section 6.

The **principle of global collaboration** is necessary for collaboration in the identification of real names across borders. This principle states that "in order to implement the principles of the Bright Internet on a global scale across borders, it is essential that Internet user countries collaborate globally in terms of communication, cooperation, execution, and reporting" (Lee et al., 2018, p. 74). Through a collaborative search, the real names of malicious attackers outside borders can be identified. However, cyberterror countries will not take part in such a collaborative effort. It is important to recall that the need for international cooperation between states is stressed by the CECC.

Overall, the above preventive security principles should not infringe on the privacy of innocent netizens. Thus, the **principle of privacy protection** is also necessary. It states, "The Bright Internet system should be technically and legally designed in consideration of protecting privacy, which may be threatened by adopting security-related principles" (Lee et al., 2018, p. 74). Assuring the auditing capacity for uncovering the footprints of illegal access of private information both technically and legally will mitigate privacy infringements. This principle may be agreed upon by the member countries of the Bright Internet, but cyberterror countries will not conform to it. Therefore, again, the Internet Peace Principles and an enforcing measure are necessary to avoid the loophole inherent within the Bright Internet principles alone.

4 Characteristics of State-led Cyberattacks

To analyze the characteristics of SLCAs in contrast to PLCAs, we present seven SLCA cases and three PLCA cases, as summarized in Tables 1 and 2 (in the Appendix), respectively. We collected the data for these cases from various open sources, such as study results from research institutions, governmental investigation results, annual reports, white papers, and the testimony of North Korean defectors.[2] We collected SLCA cases from validated sources, such as the FBI (2014); Zetter (2011); Tikk et al. (2010); the Supreme Prosecutors' Office (2015); the Korean

[2] Kim Heung-Kwang revealed North Korea's Cyber Unit' Reality, 2015-01-26, Retrieved from http://www.rfa.org/korean/weekly_program/rfa_interview/rfainterview-01262015095622.html

Ministry of Science, ICT & Future Planning (2013); and the Korean National Police Agency (2009), as listed in Row 12 of Table 1. To complete the data for the eleven factors, we investigated many additional sources, as given for each quoted data point in Table 1. We collected PLCA cases from Stefanek (2016), the Korean National Police Agency (2011), and Charette (2012), as listed in Row 12 of Table 2, and we also referenced the sources of supplementary data for each quoted data point in Table 2. We use asterisks to mark primary references, and identify other supplementary sources with the letters (a), (b) and (c). Table 3 in the Appendix also shows the attack routes, and the references are provided in the far-right column.

To characterize the cases, we compared eleven factors. These included the purposes of the attacks, the targets of the attacks, the origin country, the attack means and methods, the attack routes, and the timing, duration, preparation period, investigation period, consequences, and applicable laws. Collecting many SLCAs is very difficult because confirming the identities of attackers is not often possible given that the suspected countries typically do not admit that they harbored the attackers. It is also difficult to access the proper data sets to gain the full spectrum of the eleven factors listed above. For instance, we could not include recent cyberattack cases, such as Russia's interference in the U.S. presidential election and the DDoS attack against domain name servers in 2016, because we could not collect the entire spectrum of data. However, in the cases we used, there was sufficient nontechnical evidence that could pinpoint the originating countries.

Through this process, we selected seven typical cases from twenty major cyberattacks (see footnote) [3] as presented below.

S1) Attack on Sony Pictures Entertainment in 2014, which destroyed and leaked information as the first SLCA that targeted a particular private corporation.

S2) Stuxnet attack on Iran's nuclear facility in 2010.

S3) DDoS attack against Georgia along with a traditional military campaign in 2008.

S4) DDoS attack against Estonia in 2007, which was the first national-level cyberattack.

S5) Hacking against the Korea Hydro & Nuclear Power Co., Ltd. (KHNP) in South Korea in 2014.

S6) 3/20 Advanced Persistent Threat (APT) attack that caused major economic damage (US$867.2 million) to South Korea in 2013.

S7) 7/7 consecutive cyberattacks against the USA and South Korea in 2009.

Collecting PLCA cases is also difficult because companies that experience damage tend to hide this fact to avoid giving the impression that their IT resources are vulnerable. For comparison, we selected three typical types of PLCAs out of the many cyberattacks that occurred during the last fifteen years.

P1) The ransomware attack against Hollywood Hospital in 2016, which encrypted patient data using ransomware and demanded money to decrypt the data.

P2) The hacking and advanced persistent threat attack against SK Communications in Korea in 2011, in which information of about 35 million customers was stolen.

P3) The information-leakage attack against Coca-Cola in 2009, which stole negotiation information between Coca-Cola and the China Juice Company.

Based on these cases, we compare the commonalities and differences between the SLCAs and the PLCAs for each factor. Despite some commonalities in the basic technologies between the two types, SLCAs differ fundamentally in terms of the types of attacks, the actors, the purposes, the targets, the levels of the techniques used, the attack timing, the language, and the specific malware used. Therefore, it is necessary to establish appropriate preventive measures and countermeasures against SLCAs.

4.1 Purposes of Cyberattacks

States may use cyberattacks as a diplomatic tool with political purposes or as a national warning against potential enemies by causing social chaos and economic damage, threatening national defense and critical functions, and obtaining information about military and critical national infrastructure. This is a preferred option because it can be done without serious criticism from the international community and without fatal retaliation. As an example, in Case S1, North Korea warned the USA, the United Nations, and their allies not to play the film "The Interview" before they attacked Sony in 2014.

[3] There are a number of major SLCAs not included in the study from 2003 to 2016, such as (1) the 2016 Russian interference in the U.S. presidential election, (2) the 2015 Australian supercomputer hack, (3) the 2014 full-fledged battlefield of cyber warfare between Russia and Ukraine, (4) the 2013 6/26 DDoS attack on South Korea's government agencies, (5) the 2013 anonymous attack on Singapore in response to web-censorship regulations, (6) the 2012 Saudi Arabia Aramco attacks with the virus "Sharmoon," (7) the 2012 attacks on Iran with the computer malware Flame, (8) the 2011 Jasmine Revolution, (9) the 2010 cyberattacks on Myanmar related to the 2010 general election in Myanmar, (10) the 2010 Indian-Pakistan cyber conflict, (11) the 2008 cyberattacks on U.S. military computers, (12) the 2002-2005 Titan Rain attack to collect defense-related information in the U.S. and (13) the 2003 1/25 Slammer attack.

SLCAs also exploit the vulnerability of an enemy as an instrument for weak states or even private actors in order to inflict serious damage on more powerful states with greater traditional military power (Roscini, 2010). In Case S7 of the 7/7 DDoS attack, North Korea attacked 14 U.S. government agencies and 22 Korean government agencies (Clarke & Knake, 2010). SLCAs may occasionally be operationalized by nongovernmental organizations (Kozlowski, 2014).

On the other hand, the primary purpose of PLCAs is to gain ransom money, leak personal information, gain the upper hand in competition, show off skills, gain private benefits, and promote the attacker's personal and/or religious interests. PLCAs increasingly use ransomware to extort money in return for a solution to the encrypted data, as in the Hollywood Hospital Case P1 and the recent WannaCry attack.

4.2 Targets of Cyberattacks

The countries most frequently attacked by SLCAs are South Korea and the USA, along with others such as Iran, Georgia, and Estonia. The primary attack targets in these countries are critical infrastructure, such as energy systems, government organizations, financial institutions, and broadcasters. Attackers also target defense agencies and weapon systems, the general public, and private corporations, as demonstrated below.

4.2.1 Energy Systems

Typical targets in energy systems are nuclear power plants and electricity transmission systems, as in Case S5, in which a nuclear power plant in South Korea was attacked in 2014. An extension of nuclear power plants is nuclear weapons, which may be directly related to military systems. An example of deterring the development of a nuclear weapon was Case S2, i.e., the Stuxnet attack against an Iranian nuclear facility in 2010.

4.2.2 Governmental Organizations

In Case S4, there was "a series of massive coordinated cyberattacks on the Estonian public and private sectors" in 2007 (Rehman, 2013). These included the national assembly, governmental organizations, newspapers, broadcasters, and banks. Case S7, the 7/7 DDoS attack was conducted in July of 2009 (Digital Times, 2010) against 14 U.S. government organizations and 22 South Korean governmental organizations, including the presidential office, the Ministry of National Defense, and the National Assembly in Korea.

4.2.3 Financial Institutions and Broadcasters

All financial services and payments are closely interconnected by the Internet and dedicated networks. Thus, financial institutions are the Achilles heel of society. The attack of Case S3 occurred against Georgia in June of 2008. This DDoS attack aimed to disable banks, educational facilities, government organizations, businesses, Western press companies, and Georgian hackers' websites (Kozlowski, 2014). The DDoS attack paralyzed the banking system and ATM services and deprived the government of control over its social systems (Tikk et al., 2010). Case S6 involving the 3/20 APT attack was conducted in March of 2013 against three banks (Shinhan, Nonghyup & Jeju Banks) and three major broadcasters (KBS, MBC, & YTN) in South Korea (Jang, 2013).

4.2.4 A Particular Corporation

A SLCA against a particular corporation occurred when Sony produced the movie "The Interview." The hackers attacked Sony to damage the entire set of interior data and to leak unreleased movies.

4.2.5 General Public

SLCAs may target the general public on popular websites such as internet service providers (ISPs), portal sites, e-mail sites and the websites of public institutions. Such computer resources may be maliciously compromised. For instance, popular sites in South Korea, such as Naver, Daum, auction.co.kr, nalsee.com, conservative associations' homepages, and computer vaccine companies' homepages, were targeted by attackers, who compromised them as zombie websites during the 7/7 DDoS attack.

In contrast, PLCAs aim to attack enterprises, individual companies, government institutions, and hospitals. The targets are similar to those of SLCAs although the purposes may differ.

4.3 Origin Country

To trace the origin of SLCAs, we need to identify the origin IP address and its country. However, in the current Internet environment, attackers can easily spoof the locations of origins by compromising other computers and detouring though virtual private networks (VPNs) via third countries. According to Table 3, the attacks on the nuclear power plant in Case S5 had four originating locations in three countries: North Korea, Beijing and Shenyang in China, and Russia.

The attack on Sony Pictures in the USA originated from Chinese IPs, although they were exclusively used by North Korea. The man-in-the-middle attacks in Case S2 of the Iran Stuxnet attack required that the

attackers physically access the target. The cyberattack against the Iran nuclear facility, known as the "Olympic Game" (Sanger, 2012) attack, was operationalized in 2010, and the USA and Israel were suspected of being behind the attack (Kushner, 2013) although both governments deny involvement in the operation.

Nevertheless, five cases have used IPs in the attacker's own country, possibly because it is not easy to secretly locate an entire attacking troop in a foreign country. Even Case S5 included an IP address of its originating country.

- In Case S3 involving the cyberattacks against Georgia, the attack was conducted during the Russian engagement in the Georgian attack on South Ossetia, which has a pro-Russian government. The attack was carried out by patriotic hacktivists and attackers who used Russian government IPs, but the Russian government denied any connection to this attack.
- In Case S4, the Estonian government found that Russian computers were involved in the attack, but the Russian government denied any involvement (Ruus, 2008). In the beginning, amateur hacktivists started the attack, but during the final stage, specialists joined in the attack (Ruus, 2008).
- In Case S5 involving the hacking of a nuclear power plant in South Korea, the attack was carried out by attackers with North Korean IPs who used the malicious code that was used during other North Korean cyberattacks against South Korea.
- In Case S6 of the 3/20 APT attack, a North Korean agency (the Reconnaissance General Bureau) with North Korean IPs exploited a South Korean vaccine update program in a patch management system (Digital Times, 2010).
- Case S7 involving the 7/7 DDoS attack originated in North Korea.

According to these seven cases, the most frequent origins of cyberattacks are North Korea and Russia, and only a few countries, such as the USA, Russia, China, Iran, and North Korea are known to have the capability to conduct SLCAs (Clarke & Knake, 2010). Therefore, tracing the attacker's origin IP address is very important to identify the responsible country, although they may spoof the IP address on the current IPv4 platform. This is why the importance of implementing an unchangeable origin IP address protocol is stressed, possibly on the IPv6 platform, in the design of the Bright Internet.

On the other hand, PLCAs are conducted by private actors; accordingly, the origin IPs can be any country, although the three PLCAs cases in Table 3 either originated in China or were unidentified. It is essential for ransomware attackers such as those in Case P1 to conceal their origins and use Bitcoin. In this regard, the anonymous cryptocurrency has the capacity to facilitate criminal activities.

4.4 Attack Means and Methods

There are four main attack means and methods[4] used for cyberattacks: hacking, the spreading of malicious code, distributed denial of service (DDoS), and combined attacks (Kang et al., 2010). According to 2017 Cost of Cyber Crime Study (Ponemon Institute, 2017), the percentages of cyberattacks in the four categories experienced by companies in 2017 are as follows:

(a) Hacking: web-based attacks (67%)

(b) DDoS: DoS (53%), botnets (63%)

(c) Spreading of malicious: malware (98%), malicious code (58%), ransomware (27%);

(d) Combined attacks: phishing and social engineering (69%), malicious insiders (40%), and stolen devices (43%).

According to the seven SLCA cases in Table 1, the following attack methods are used:

(a) Hacking attacks including APT attacks, homepage defacements, web server attacks and zero-day attacks: Relevant cases are S1, S2, S3, S4, S5, S6, and S7 in Table 1 and P1, P2 and P3 in Table 2.

(b) DDoS: Cases S3, S4, and S7.

(c) Spreading of malware: Cases S1, S2, S5, S6, and S7 and P1, P2 and P3.

(d) Combined attacks: Cases S5 and P3.

According to the case analysis, we observe that SLCAs and PLCAs use different attack means and methods. SLCAs use more sophisticated and proven technologies than PLCAs. For example, cyberweapons such as Stuxnet are developed in secret over a long period of time with better financial support (Chang, 2012, p. 6) and are carried out as combined attacks. The same unique malware, such as Kimsuky, was used repeatedly during SLCAs by the same country. Typical means and methods of SLCAs are new, combined, and advanced technologies, such as zero-day exploitations of new computer

[4] Methods are cybertactics, techniques, and procedures, and means refer to cyberweapons and their associated cybersystems according to the Tallinn Manual of 2017.

vulnerabilities, as analyzed in Table 1. The APT attack was a typical tailored attack against a specific target involving long-term preparation by employing a combination of methods, such as social engineering, zero-day attacks, rootkits, and sometimes DDoS attacks as well. For the compromise and spreading of malicious code, advanced account hijacking and worms such as Kimsuky, Wiper, Dark Seoul and MyDoom were used in the cases involving South Korea, and Destover was used in the U.S. attacks. In contrast, PLCAs mainly use ransomware, phishing, keyloggers, and malware for swindling purposes and to obtain information.

4.5 Attack Routes

SLCAs route through many ISPs and VPNs in both origin and destination countries. As described in Table 3, all SLCAs and P2 exploited multiple routes from origins to destinations using direct and indirect courses through a third country, mostly through China or the USA, but also through Japan, Thailand, and the Netherlands. Thus, implementing the deliverer responsibility principle and global collaboration on the Bright Internet is very important in order to prevent detoured attacks. SLCAs detoured through many more countries to conceal the attackers' identity than did PLCAs, as shown in Tables 1 and 2.

4.6 Attack Timing

Cyberattacks can be conducted at any time and from anywhere. However, SLCAs are usually initiated when important national issues arise and when sending a warning message is deemed necessary. North Korea conducted a cyberattack as a warning message to the USA on Independence Day. Cyberattacks may be operationalized as a weapon immediately before and during a traditional war, as in Case S3. Hacktivists may attack another group when a social conflict occurs within a country or between countries. On the other hand, criminal PLCAs are conducted at a time of interest for the private attacker.

4.7 Attack Duration

The attack duration of SLCAs is usually a few months, while PLCAs tend to take only a few days. However, the duration depends upon the incident. SLCA cases in this study lasted on average 74 days per attack, while PLCA cases lasted only 13 days on average, because SLCAs tend to be more complicated, staged, and large-scale attacks. For instance, the hacking of the nuclear power plant in Case S5 continued for four months. However, the average duration of PLCA cases chosen for this study appears to be longer than typical, judging from the statistics in the consequences section.

4.8 Preparation Period

SLCAs usually require longer preparation than PLCAs because SLCAs tend to be more elaborate and large-scale events than PLCAs. SLCAs usually take at least a few months or even a few years depending upon the time of need and time necessary to develop and penetrate the target system. The well-timed attacks against Sony Pictures right before the release of the movie shortened the preparation time to three months, while the Stuxnet attack against Iran took many years to prepare.

4.9 Investigation Period

According to the cases in Tables 1 and 2, an investigation of a SLCA took 57 days on average from the occurrence of the incident to its first announcement of the result based on the analysis of S1, S5, S6, and S7. On the other hand, PLCAs had shorter periods than SLCAs due to their relatively small scale and limited scope.

4.10 Consequences

SLCAs usually damaged the target by destroying PCs, main boot records, hard disks, and data. It is neither easy to investigate the consequences of SLCA incidents nor easy to calculate the damage caused. SLCAs typically cause major economic damage (e.g., US$867.2 million in the 3/20 APT attack in Case S6) and social disorder, while also threatening national security. The average amount of damage in four cases of S1, S4, S6, and S7 in Table 1 was US$260 million per attack. According to the Ponemon Institute (2015), cyberattacks in general caused average damages of US$78,000 per attack. [5] Thus, we find that the damage caused by SLCAs (US$260 millions in Table 1) is about 3,000 times greater than that caused by PLCAs.

4.11 Applicable Laws

Wrongful conduct by SLCAs can be regarded as national illegal activity, violations of international duty, and infringements on other countries' national interests. Currently, the Mutual Legal Assistance Treaty may be used as a platform to support international collaboration, but it is not entirely effective. Hence, when SLCAs cause physical damage and destruction, the conduct should be governed by the spirit of the UN Charter, the

[5] According to a survey of 252 companies in seven countries, each company encountered an average of 1.9 cyberattacks per week, costing an average of US$77,935 per attack, and US$7.7 million per year. The highest annual average was US$15.42 million in the U.S., while the lowest was US$2.37 million in Russia (Ponemon Institute, 2015).

Responsibility of States, the Geneva Convention and the Hague Convention. However, these conventions have not yet been extended to cyberspace. In particular, when SLCAs result in the breaching of an international obligation according to the UN Charter and the Tallinn Manual, the attacked country can invoke legitimate countermeasures. If incidents involve the use of force, they should be reviewed by the United Nations Security Council. In contrast, PLCAs can be appropriately regulated according to the national criminal laws of each country.

5 Application of the Bright Internet against State-Led Cyberattacks

In this section, we elaborate on the characteristics of SLCAs vis-à-vis PLCAs from the perspective of the Bright Internet principles, as depicted in Figure 2. For this purpose, we call the global governance body that would enforce the Bright Internet principles the Bright Internet Global Organization (BIGO). We review how SLCAs that originate in BIGO member countries could be transparently identified and thus prevented, whereas those from BIGO nonmember countries (non-BIGO countries) would not be preventable. Concerning the PLCAs from non-BIGO countries, it is uncertain whether these countries would cooperate to identify private origins. This is why the complementary measure of the Internet Peace Principles is necessary, as we discuss in Section 6.

To deploy the five principles of the Bright Internet, all countries who use the Internet should accept the principles and install the appropriate systems necessary to implement the principles with due diligence. However, the adoption of the Bright Internet does not mean that the current Internet will be replaced entirely. The Bright Internet may be selectively adopted for essential applications. We now turn to an analysis of SLCAs in the context of Bright Internet principles.

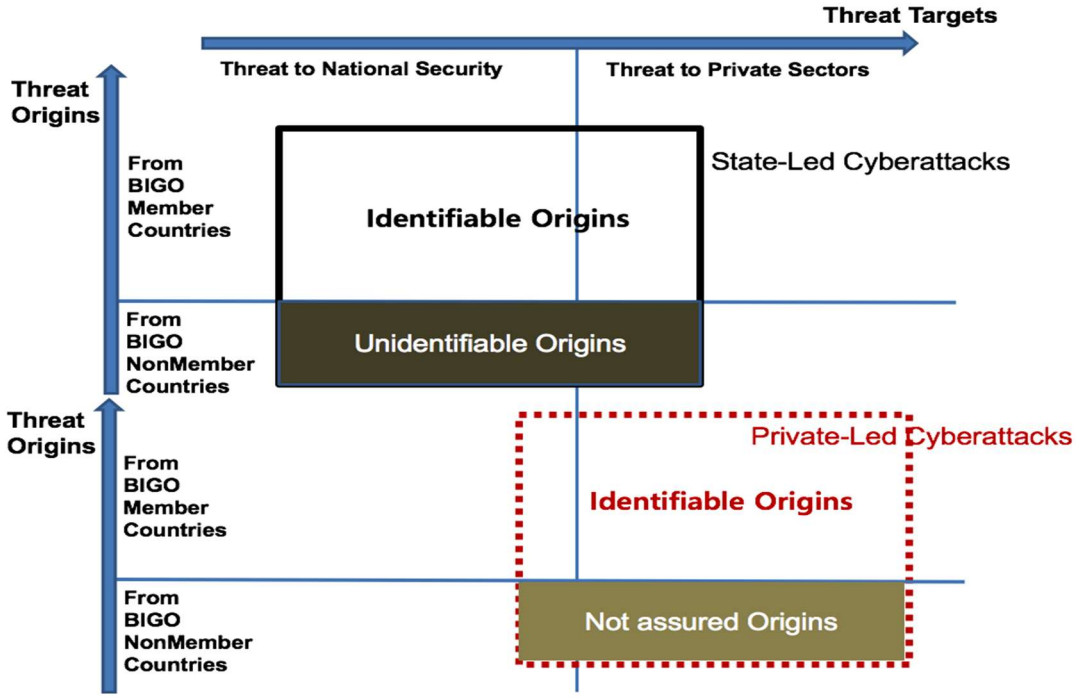

Figure 2. Origins and Targets of Cybersecurity Threats

5.1 Origin Responsibility

In our case analysis, we determined that the traced origin IP address is the major form of evidence used to identify responsible countries. Note that most attacking countries use their own resources within their own countries, although they sometimes also detour the origin of certain attacks. However, no country under suspicion would likely agree with the claims of a country under attack. Therefore, it is clear that the origin IP address alone is not sufficient to determine legal responsibility in the context of international legal systems. Accordingly, we must consider other supplementary factors to identify the originator of SLCAs, such as attack methods, attack timing, and circumstantial evidence, such as the purpose and consequences of the attack, as described below.

Origin IP Address as Evidence: The origin IP address tends to be located within the attacker's own country. To ensure the traceability of origin IP

addresses, it will be useful to adopt the IPv6 protocol[6] (Nygren, 2015), since this platform can support antispoofing (Geers, 2011).

Attack Means and Methods as Evidence: Some SLCAs may use the same means and methods regardless of the location of the originating country. For instance, the attacks of Case S5 that originated in China used the same malware "Kimsuky" that was used by North Korean hackers. Special language characters and the repeated use of similar malware can serve as evidence of the origin country. However, there is a small risk that attackers may maliciously use the malware of another attacker to disguise the origin.

Attack Timing as Evidence: The attack timing and consequences are closely interrelated in most cases. Therefore, the attack timing can serve as supportive evidence, but is not definitive. In Case 5, the attacks from North Korea and China occurred at the same time.

Circumstantial Evidence: The surmised purpose and its consequences can serve as circumstantial evidence. These factors include the prevailing political environment, the top leader's expressed intention to retaliate, the nature of the target, and whether the operation portends the future use of military force (Schmitt, 2017).

As shown above, a combination of the origin IP address, the attack methods, timing, and circumstantial evidence are useful for the identification of origins, although they cannot be taken as absolute evidence. One reason that countries currently use cyberattacks to attack their enemies is to avoid serious legal repercussions. Therefore, it will be necessary to increase the number of BIGO member countries abiding with the Internet Peace Principles.

5.2 Deliverer Responsibility

5.2.1 Detoured Attacks: Using ISPs

If an attacker sends malware through VPNs, The onion router (TOR) networks, or proxy servers, platform managers such as ISPs should be able to detect suspicious packets. BIGO countries could help each other prevent detoured attacks through their own countries. However, some states may not legally allow surveillance of their ISPs. CECC member countries, China, Russia and others allow such surveillance. However, the USA and South Korea do not allow this type of surveillance in order to protect the privacy of Internet users. However, the governments of these countries recently passed bills allowing surveillance in order to prevent terrorism.[7] It has become a global trend for most governments to require ISPs to assume deliverer responsibility for national security.

5.2.2 Wide Range of Victims: Using Zombie PCs

Compromised (zombie) computers in detoured countries themselves may not have even recognized that they were compromised, but they can be manipulated by malicious state-led hackers, as shown in Cases S3, S4, and S7. Therefore, detection and prevention technologies for routers and PCs must necessarily be developed so that they can avoid becoming compromised. Deliverer responsibility will be also a channel for identifying the original attacker behind zombie PCs. Through the international collaboration of deliverer responsibility, cyberattacks outside of an attacker's country could be detected by associated BIGO member countries. However, detoured cyberattacks through nonmember countries would remain undetectable. Therefore, it is essential that every country becomes a BIGO member and exercises jurisdiction over its own cyberpersonnel, equipment and facilities in order to avoid being targeted as a detoured cyberterror haven.

5.3 Identifiable Anonymity

Under current Internet protocol, the identification of the real names of anonymous hackers is not possible because the origin servers are not required to have the capacity to identify real names. A case study demonstrated that hostile countries use their own IPs, but they refused to cooperate in identifying the responsible originators. Accordingly, the investigation of real originators would not be possible for non-BIGO countries.

To allow for the freedom of anonymous expression for innocent netizens, the Bright Internet adopts the principle of identifiable anonymity. This principle can be implemented by adopting two layers of names: real names and optional pseudo-names (Lee et al., 2018). This principle could be implemented by encouraging a voluntary real-name IP account in order to gain the benefit of trustworthiness, as is done in the credit card industry. In case users need a space for anonymous expression, they could gain an additional application-level account, such as a social network or portal site, which would limit other types of computation that could be abused for hacking. Thus, within a BIGO member country, the real names of cyberattackers could be identified if a valid search warrant were

[6] As of June 24, 2017, 37.3% of Internet users in Belgium adopted IPv6, 24.8% in the U.S., 12.0% in Japan, 2.0% in South Korea, and 0.8% in China (Akamai, 2017).

[7] South Korea: Anti-Terrorism Act of 2016; U.S.: Cybersecurity Information Sharing Act of 2015.

issued, but again this would not be possible with non-BIGO countries.

5.4 Global Collaboration

A timely investigation of a SLCA requires prompt and effective international investigative cooperation. Nevertheless, cooperation with hostile countries will not be easy for both technical and political reasons. Computer investigations require sophisticated investigational skills and processes as well as the preservation of evidence on computers and in related facilities. In Case S4, Russia refused a joint investigation proposed by Estonia in 2007, and North Korea has consistently denied any involvement in cyberattacks against South Korea. The different legal systems in different countries can also delay such investigations. In the Philippines, the virus "I Love You" was spread globally via e-mail in 2000, causing US$8.7 billion of damage, but the culprit could not be indicted because there was no applicable Philippine law in place (Deflem & Shutt, 2006).

Under the BIGO Agreement, the real names of cyberattackers could be identified among BIGO member countries even across borders if an internationally valid search warrant were issued, but again this would not be possible in non-BIGO countries. A standard global legal and technical framework for collaborative searches is essential to protect against detoured cyberattacks from third countries. However, the prevention of attacks from non-BIGO countries would require another deterrence measure.

5.5 Privacy Protection

The privacy of innocent netizens on the Bright Internet will be maintained by distinguishing innocent netizens from malevolent criminals and terrorists. Real names will be required only when a valid authority detects criminal evidence. To protect the privacy of innocent citizens, transparent audits by a trusted third party could be adopted. Privacy assurance technologies that log the footprints of illegal access to prohibited private data would also be necessary (Agre & Rotenberg, 1998). Emerging blockchain technology may be adopted to implement this principle. When a real name is stored for the implementation of identifiable anonymity, an adequate encryption method could be used so that private information would not be leaked and this information would not be released without a valid search warrant. These issues will require substantial research to fulfill privacy aims while maintaining operational efficiency. However, there would be no guarantees of privacy protection for netizens in non-BIGO countries.

6 Internet Peace Principles

In the previous section, we noted that Bright Internet principles alone cannot prevent malicious SLCAs from non-BIGO countries. Thus, it is necessary to derive a complementary measure in the form of the **Internet Peace Principles,** which we define here as follows: The (Bright) Internet is built to enhance peaceful communication and understanding between all global inhabitants and thus should neither be used as a weapon for attacking other countries nor as a means of detoured malicious attacks. To derive the Internet Peace Principles, we adopted three types of reasoning: the extension of the physical convention to cyberspace, the expansion of international cyber security conventions to global member countries, and analogical international norms. By integrating the derived principles with the specific principles of the Bright Internet, we derived the Internet Peace Principles, as depicted in Figure 3. We removed redundant statements and resolved conflicts between statements by prioritizing their importance.

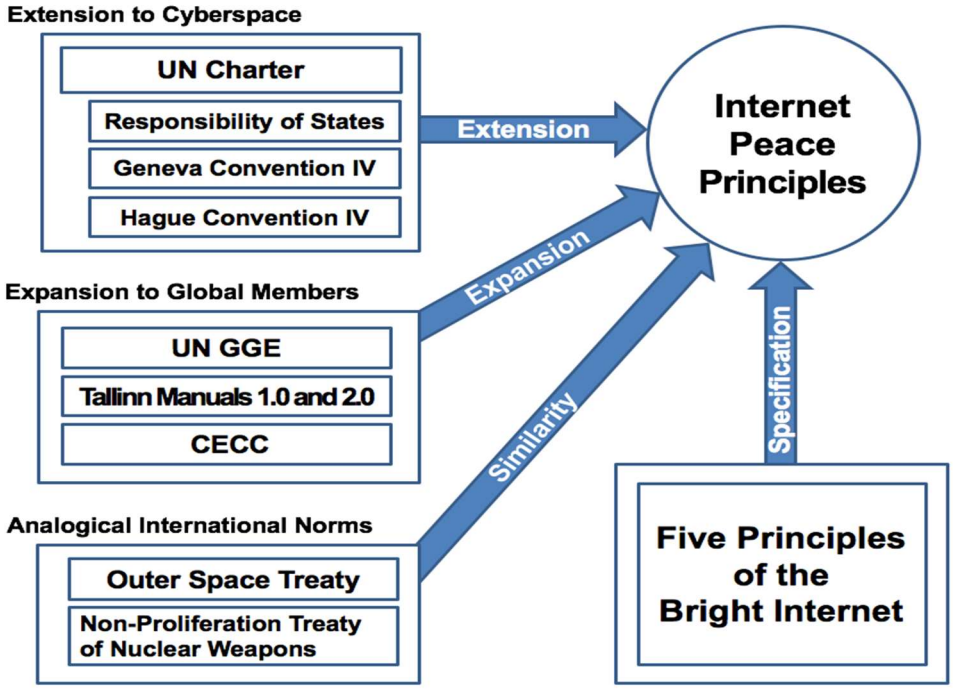

Figure 3. Methodologies for Deriving the Internet Peace Principles

6.1 Extension to the Cyberspace Approach

We derive the spirit of the Internet Peace Principles from the existing international conventions established for physical space, which should be extended to cyberspace. The most relevant cases include the UN Charter, the Responsibility of States, the Geneva Convention, and the Hague Convention. Based on these established agreements, we can derive the specific concept of the Internet Peace Principles for the safety of cyberspace. The extended statements are italicized below.

- The **UN Charter** can be extended to cyberspace, as follows: Maintain international peace and security *in cyberspace*; take effective collective measures for the prevention and removal of *cyber*security threats and for the suppression of *cyber*attacks; prohibit the use of *SLCAs* except for legitimate self-defense and countermeasure purposes; the UN Security Council may take . . . demonstrations, blockade and other operations by air, sea, land and *cyber* forces of members of the United Nations.
- The **Responsibility of States for Internationally Wrongful Acts** can be extended to cyberspace as follows: The offending state should be held responsible for its *cyber*attacks, and the attacked state is entitled to request compensation and take legitimate countermeasures.
- **The Geneva Convention IV** can be extended to cyberspace as follows: Prohibit preemptive *cyber*attacks against innocent civilians and their facilities *in cyberspace*.
- The **Hague Convention IV** can be extended to cyberspace as follows: Regulate the unlimited use of means and methods of *cyberwarfare*.

6.2 Expansion of International Cyber Security Conventions to Global Members

The UN GGE and the Tallinn Manual specify the cybersecurity aspects of each participating country, while the CECC does so for each participating organization and individual. However, the UN GGE is still in the recommendation stage, and the Tallinn Manual is a nonbinding document. The CECC has only 55 member countries as of June 27, 2017.[8] Therefore, it will be necessary to build a comprehensive global convention for cybersecurity which can also be merged with the specific principles of the Bright Internet, as italicized below.

- The **UN GGE** recommendations can be adopted for the Internet Peace Principles as: States should not knowingly allow their territories to be used for internationally wrongful acts using ICTs *including the breaching of the principles of*

[8] Retrieved June 27, 2017, from http://www.coe.int/en/web/conventions/full-list/conventions/treaty/185/signatures?p_auth=Z9kSyEKq

origin responsibility, deliverer responsibility, and privacy protection. States should not *cyber*attack the critical infrastructure of other states. States should cooperate to exchange information to assist each other and to prosecute terrorist and criminal use of the Internet *including the identification of real names upon the presentation of a valid search warrant.*

- **The Tallinn Manual** provides rules about state actors' responsibilities as well as their jurisdictions and countermeasures during cyberconflicts. Rules 6-7 (due diligence) specify the state's control responsibility of cyber-infrastructure that correspond with the chain of origin responsibility. Rule 14 states that a state shall bear an international legal responsibility for a cyberoperation attributable to it. This corresponds to the extension of the responsibility of a state to cyberspace. The manual mentions "jus ad bellum," the right to go to war, by specifying the conditions of a cyberoperation that justifies a response in the form of self-defense (rule 71) and countermeasures (rule 20), which corresponds to the UN Charter. Rules 94 and 99 invoke "jus in bello," a humanitarian law that regulates conduct during armed conflicts by specifying that launching cyberattacks against civilians or civilian objects is unlawful, and that corresponds to the extension of the Geneva Convention to cyberspace. As presented above, these rules in the manual overlap with those of the UN GGE and/or the cyberextension of existing conventions.

  Rule 109 is a unique statement *which extends the prohibition to nuclear power plants.* However, the rules about collective cybersanctions and cyberblockades are contradictory with regard to the cyberextension of the UN Charter. In rule 144, *collective cybersanctions* are prohibited, although they are essential for the implementation of deliverer responsibility. In rule 128, there is debate about rules governing cyberblockades as to whether they should be characterized as *lex lata* (referring to current law) or *lex ferenda* (referring to future law) considering the level of technical feasibility. Because cyberblockades *of cyberterror countries or zones are essential* for the execution of preventive security by deliverer responsibility, we regard a *cyberblockade as lex lata* in the Internet Peace Principles.

- The **CECC** mainly describes PLCAs and does not provide measures against SLCAs. However, we introduce the CECC here because BIGO will need to consider not only SLCA aspects but also those associated with PLCAs.

6.3 Analogical International Norms Approach

The analogical international norms are referred to as a means of deriving the concept of the Internet Peace Principles from similar international norms for similar problems. For this purpose, we refer to two analogous treaties: the Outer Space Treaty and the Non-Proliferation Treaty of Nuclear Weapons (NPT).

- Since outer space is regarded as the fourth domain, the Internet is regarded as the fifth domain after land, sea, and air. Hence, the Internet Peace Principles should be adopted in similarity with the Outer Space Treaty. The Internet space should be used for humankind and not as a type of weapon.

- From the NPT's vision of "Atoms for Peace," we can derive the notion of the "Internet for Peace." For research on and the verification of the Bright Internet and Internet Peace Principles, there needs to be a forum for agreement like the Bright Internet Global Summit (BIG Summit). For the implementation of the principles, we need an international body like BIGO, just as the IAEA is necessary for the implementation of the NPT.

6.4 Ten Internet Peace Principles

Based on the above analysis, we derive the following ten statements of the Internet Peace Principles:

- **Maintenance of Peace:** All states should maintain international peace and security in cyberspace.

- **Prevention of Misuse:** States should not knowingly allow their territories to be used for internationally wrongful acts using ICTs, including breaching the principles of origin responsibility, deliverer responsibility, and privacy protection.

- **Protection of Critical Infrastructure and Civilians:** States should prohibit preemptive cyberattacks against critical infrastructure, including nuclear power plants and innocent civilians and their facilities in cyberspace.

- **Conditions of Legitimate Cyberattacks:** States should prohibit the use of SLCAs except for in legitimate self-defense and for countermeasure purposes.

- **Regulation of Means and Methods:** States should regulate the unlimited use of the methods and means of cyberwarfare.

- **International Cooperative Search**: States should cooperate to exchange information, including the identification of real names upon

the issuance of a valid search warrant, in order to assist each other and to prosecute terrorists and the criminal use of the Internet.

- **Accountability:** The offending state should be held responsible for the consequences of its cyberattacks.
- **Entitlement for Compensation and Self-Defense:** An attacked state is entitled to request compensation and to undertake legitimate self-defense and countermeasures.
- **Enforcement Mechanism:** The UN may take effective collective measures for the prevention and removal of cybersecurity threats and for the suppression of cyberattacks and may consider cyberblockades.
- **Global Governance:** To ensure cooperation with the enforcement mechanisms, a forum like the BIG Summit and a governance body like the BIGO are necessary in order to conduct research and realize the implementation and verification of the Bright Internet and Internet Peace Principles.

However, not all countries may wish to become members of a BIGO-like organization. This is why a sanction mechanism against noncompliant cyberterror countries will be necessary.

7 Concluding Remarks and Summary

Cyberterrorism and cybercrimes on the Internet have become more sophisticated and intelligent, causing major social and economic damage, and threatening the foundation of national security and sustainable growth. In particular, as hostile countries conceive of acts of cyberterrorism more systematically, critical national infrastructure becomes more vulnerable to catastrophic disasters. However, the current individualized defensive security systems cannot identify malicious originators and prevent them from such state-led cyberattacks. Moreover, in the defense of cyberspace, there is no clear distinction between military and civilian frontiers. Nevertheless, there are currently no effective global governance mechanisms that can mitigate the threat of state-led cyberattacks.

As the foundation of a preventive security solution against the risk of state-led cyberattacks, the originators of cyberterrorism, or at least the originating country, should be identifiable. To achieve this goal, the Bright Internet initiative has proposed five principles: origin responsibility, deliverer responsibility, identifiable anonymity, global collaboration, and privacy protection. By implementing these principles, the Bright Internet intends to achieve the balanced goals of preventing cybersecurity and guaranteeing the freedom of expression and the privacy of innocent netizens. Achieving these goals simultaneously requires a holistic design of protocols, technologies, national legislation, and global collaboration and implementation with support from internationally agreeable organizations. For this purpose, we design and propose the Bright Internet Global Organization in this research.

According to our vision, as long as BIGO member countries conform to the principles of the Bright Internet, transparency of the malicious origins of cyberattacks can be maintained, and thus the motivation of state-led cyberattacks will be deterred. However, cyberterror countries will likely not voluntarily accept BIGO membership. Therefore, it will be necessary to enforce the participation of all Internet-connected countries through a new international law, like the Internet Peace Principles that we propose in this paper. The Internet Peace Principles prohibit the use of the Internet as a weapon for attacking other countries or as a means of detoured malicious attacks. To derive the Internet Peace Principles, we adopted three approaches: the extension of the physical convention to cyberspace, the expansion of international cyber security conventions to global member countries, and analogical international norms.

For the extension of the traditional convention to cyberspace, we adopted the United Nations Charter, the Responsibility of States for Internationally Wrongful Acts, the Geneva Convention, and the Hague Convention. For the expansion of international cybersecurity conventions to global member countries, we adopted the UN Group of Governmental Experts, the Tallinn Manual created through an invitation by the NATO Cooperative Cyber Defence Center of Excellence, and the Council of Europe Convention on Cybercrime. Finally, for the derivation of similar principles from analogous international norms, we adopted the Outer Space Treaty, and the NPT. By integrating the concepts from these sources in a consistent manner, we derive the statements pertaining to the Internet Peace Principles, as described in Section 6.

Owing to the deterrence power of the Bright Internet and the Internet Peace Principles, we believe that the paradigm of preventive cybersecurity can be established and that the motivations behind cyberattacks can be drastically ameliorated. Therefore, all countries should come together to establish this type of framework and a Bright Internet global governance organization. To facilitate talks toward this end, the BIG Summit was held for the first time in Seoul in December of 2017 as a pre-ICIS conference of the Association for Information Systems. In 2018, two meetings will be held at

Tsinghua University and Stanford University respectively. By establishing the Bright Internet Global Organization with an appropriate sanction mechanism, national security and human society would be able to prevent the catastrophic threat of state-led cyberattacks.

The research on preventing state-led cyberattacks through the Bright Internet and the Internet Peace Principles is in its infancy, but we believe it has the potential to have a major effect and represents a significant research opportunity in the effort to make the next generation of the Internet safer and more peaceful. We can foresee a major research opportunity in the area of security policies for preventive information systems domestically and internationally. This opens new research opportunities in areas of design science and behavioral science. A wider discussion on this topic can also be found in the paper, "Design and Validation of the Bright Internet," found in part 1 of this special issue on the Bright ICT (Lee et al., 2018).

Acknowledgements

This paper was based on an unpublished thesis written by Young Yung Shin, which was supervised by Jae Kyu Lee and Myungchul Kim, and entitled "Preventive Measures against State-led Cyberattacks by the Bright Internet and Internet Peace Principles" at the Korea Advanced Institute of Science and Technology (KAIST) in 2017. This research was supported by grants from the KAIST Bright Internet Research Center, the Climate Change Research Hub of KAIST (Grant No. N011170062), the Institute of Information and Communication Technology Promotion, and the National Research Foundation of Korea funded by the Korean government (No. 2017R1A2B4005865). This research was also supported by Xi'an Jiaotong University in China and Heinz College of Carnegie Mellon University in the USA.

References

Agre, P. E. & Rotenberg, M. (1998). *Technology and privacy: The new landscape*. Cambridge, MA: Massachusetts Institute of Technology Press.

AhnLab. (2010, October). New Paradigm in Malware, Stuxnet. *Online Security Magazine Monthly Ahn* [in Korean]. Retrieved from http://www.ahnlab.com/kr/site/securityinfo/secunews/secuNewsView.do?cmd=print&seq=16852&menu_dist=2

AhnLab. (2013, April). Threat analysis: Malicious code detail analysis according to types problems in computer booting relating to 3/20 APT attack, *Online Security Magazine Monthly Ahn* [in Korean]. Retrieved from http://www.ahnlab.com/kr/site/securityinfo/secunews/secuNewsView.do?cmd=print&seq=20767&menu_dist=3

Akamai. (2016). IPv6 adoption visualization. Retrieved from https://www.stateoftheinternet.com/trends-visualizations-ipv6-adoption-ipv4-exhaustion-global-heat-map-network-country-growth-data.html

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: A multimedia empirical examination of home computer user security behavioral intentions. *MIS Quarterly*, *34*(3), 613-643.

Azzopardi, M. (2013). The Tallinn Manual on the international law applicable to cyber warfare: A brief introduction on its treatment of jus ad bellum norms, *ELSA Malta Law Review,3*, 174-184.

BBC News. (2012). Coca-Cola "targeted" by China in hack ahead of acquisition attempt. *BBC News.* Retrieved from http://www.bbc.com/news/technology-20204671

Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly*, *35*(4), 1017-1042.

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, *34*(3), 523-548.

CECC. (2001). *Council of European convention on cybercrime*. Retrieved from http://www.europarl.europa.eu/meetdocs/2014_2019/documents/libe/dv/7_conv_budapest_/7_conv_budapest_en.pdf

Chae, S. W. (2010, July). One year after the 7/7 DDoS attack: What's the secret of number 7? [in Korean]. *Digital Daily*. Retrieved from http://www.ddaily.co.kr/news/article.html?no=65423

Chang, N. S. (2012). Cyber weapons & international security, *Jeju Peace Institute Policy Forum*, *2012, 19*. Retrieved from http://www.jpi.or.kr/kor/regular/policy_view.sky?code=assay&id=4735

Charette, R. N. (2012, November). Coca-cola disrespects its investors by not telling them about massive 2009 computer breach, *IEEE Spectrum.* Retrieved from http://spectrum.ieee.org/riskfactor/telecom/security/cocacola-neglects-to-tell-investors-about-massive-2009-breach

Chen, Y., & Zahedi, F. M. (2016). Individual's Internet security perceptions and behaviors: Polycontextual contrasts between the United States and China. *MIS Quarterly, 40*(1), 205-222.

Choe, S. H., & John, Markoff. (2009, July 8). Cyberattacks jam government and commercial web sites in U.S. and South Korea. *Xinhuanet.* Retrieved from http://www.nytimes.com/2009/07/09/technology/09cyber.html

Clarke, R. A. & Knake, R. K. (2010). *Cyber War*. New York, NY: Harper Collins.

Csanyi, E. (2011, August). How Stuxnet (PLC virus) spreads: Part 3. Electrical Engineering Portal. Retrieved from http://electrical-engineering-portal.com/how-stuxnet-plc-virus-spreads-part-3

D'Arcy, J., Hovav, A., & Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research, 20*(1), 79-98

Daily Mail. (2014, December 12). "Not only you but your family will be in danger": Sony hackers flash stark warning across employees' screens as they boast that the damage so far is just tip of the iceberg. *Daily Mail.* Retrieved from http://www.dailymail.co.uk/news/article-2871450/Sony-hackers-flashed-disturbing-new-warning-company-s-computers.html.

Deflem, M. & Shutt, J. E. (2006). Law Enforcement and Computer Security Threats and Measures. In H. Bidgoli, H. (Ed.), *Handbook of*

information security: Vol. 1. Information warfare; social, legal, and international issues; and security foundations. Hoboken, NJ: John Wiley & Sons. Retrieved from http://deflem.blogspot.com/2006/08/law-enforcement-and-computer-security.html

Digital Times. (2010, July 5). DDoS attack, multiple simultaneous raid . . . reaches several tens of billion won. [in Korean]. *Digital Times.* Retrieved from http://www.dt.co.kr/contents.html?article_no=2010070602011260600002.

Emerging Technology Research Center. (2009). Cyberterrorism and IT Korea under the current situation [in Korean]. Retrieved from http://www.etrc.co.kr/bbs/bbs.php?table=research&query=view&uid=17

Fleck, D. (2013). Searching for international rules applicable to cyber warfare: A critical first assessment of the new Tallinn Manual, *Journal of Conflict Security Law*, *18*(2), 331-351.

Gallager, S. (2014). Sony Pictures attackers demand: Stop the terrorist film! *Arstechnica*. Retrieved from http://arstechnica.com/security/2014/12/sony-pictures-attackers-demand-stop-the-terrorist-film/

Geers, K. (2011). *Strategic cyber security*. Retrieved from https://www.defcon.org/images/defcon-19/dc-19-presentations/Geers/DEFCON-19-Geers-Strategic-Cyber-Security-WP.pdf

Gil, M. G. (2014, December). Red alert team releases hacking analysis report for Sony Pictures internal networks [in Korean]. *DailySecu.* Retrieved from http://www.dailysecu.com/news_view.php?article_id=8386

Gill, T. G., & Hevner, A. R. (2013). A fitness-utility model for design science research. *ACM Transactions on Management Information Systems, 4*(2), 5:1-24.

Hague Convention IV. (1907). Convention IV respecting the Laws and Customs of War on Land and its annex: Regulations concerning the Laws and Customs of War on Land. The Hague, 18 October 1907. Retrieved from https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Treaty.xsp?action=openDocument&documentId=4D47F92DF3966A7EC12563CD002D6788

Herath, T., & Rao, H. R. (2009a). Protection motivation and deterrence: A framework for security policy compliance in organizations. *European Journal of Information Systems, 18*(2), 106-125.

Herath, T., & Rao, H. R. (2009b). Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. *Decision Support Systems, 47*(2), 154-165.

Hesseldahl, A. (2015, January 20). Here's What Helped Sony's Hackers Break In: Zero-Day Vulnerability. *Recode*. Retrieved from http://www.recode.net/2015/1/20/11557888/heres-what-helped-sonys-hackers-break-in-zero-day-vulnerability

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design science in Information Systems research," *MIS Quarterly 28*(1), 75-104.

Hruska, J. (2015, March 11). Windows PCs vulnerable to Stuxnet attack - five years after patch. *Extremetech* [Weblog post] Retrieved from http://www.extremetech.com/computing/200898-windows-pcs-vulnerable-to-stuxnet-attack-five-years-after-patches

Hsu, J. S. C., Shih, S. P., Hung, Y. W., & Lowry, P. B. (2015). The role of extra-role behaviors and social controls in information security policy effectiveness. *Information Systems Research, 26(*2), 282-300.

IAEA. (2016). International Atomic Energy Agency. Retrieved from https://www.iaea.org/sites/default/files/publications/documents/infcircs/1970/infcirc140.pdf

Infosec institute. (2016). Ransomware Case Studies: Hollywood Presbyterian And The Ottawa Hospital. Retrieved from http://resources.infosecinstitute.com/category/healthcare-information-security/healthcare-attack-statistics-and-case-studies/ransomware-case-studies-hollywood-presbyterian-and-the-ottawa-hospital/

International Committee of the Red Cross (ICRC), Geneva Convention relative to the protection of civilian persons in time of war. Retrieved from http://www.refworld.org/docid/3ae6b36d2.html

International Law Commission. (2001). Draft articles on responsibility of states for internationally wrongful acts. Retrieved from http://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf

Ireland, N. (1967). Treaty on principles governing the activities of states in the exploration and use of outer space, including the moon and other celestial bodies. Retrieved from http://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html

Jang, J. S. (2013, March 20). (4th LD) Several broadcasters, banks suffer suspected cyber attacks. *Yonhap News Agency*. Retrieved from http://english.yonhapnews.co.kr/national/2013/03/20/40/0301000000AEN20130320008051315F.HTML

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: an empirical study. *MIS Quarterly, 34*(3), 549-566.

Joint Chiefs of Staff, United States Department of Defense (2009). *Joint publication 1-02: Department of Defense dictionary of military and associated terms.* Retrieved from http://jitc.fhu.disa.mil/jitc_dri/pdfs/jp1_02.pdf

Kang, S. K., Yoon, H. S., Park, Y. W., Kim, M. H., Kwon, H. Y., Kim, D. S., & Kim, G. B. (2010). A Study on the Construction of Efficient System for Safe Cyberspace. Korea, Korean Institute Criminology. Retrieved from https://www.kic.re.kr/pubdata/public/Read.jsp?paramNttID=4362&paramPage=30

Kim, H. S. (2013, December 8). The World is On Cyber War [in Korean]. *Newskorea*. Retrieved from http://www.newskorea.com/bbs/board.php?bo_table=weekly_issue&wr_id=584

Kim, K. A. (2015, March 4). Recent Security Threat, Hacking on Router and Content Delivery Network. Retrieved from http://www.hauri.co.kr/information/news_view.html?intSeq=6875&page=1

Korean National Police agency. (2009, July 28). Multi-Dimensional DDoS Attack Configuration Confirmed Using Attack Order Confirmed as Overseas Servers and Malicious Code as Domestic Servers [in Korean]. Retrieved from http://www.police.go.kr/portal/bbs/view.do?nttId=7560&bbsId=B0000011&menuNo=200067&delCode=0

Korean National Police Agency. (2011, August 11). Investigation Authority is Looking into Details of Information Leakage of SK Coms' Membership and Tracing the Hacker [in Korean]. Retrieved from http://police.go.kr/m/bbs/view.do;jsessionid=qxk7px8VxqtQjiJP6QHVkhRuxLsa4CTuAzqb1czny1a22wayOGBaRaaGiah9mHrz.websvr02_servlet_engine1?nttId=8102&bbsId=B0000011&menuNo=2600013&delCode=0

Kozlowski, A. (2014). Comparative Analysis of Cyberattacks On Estonia, Georgia & Kyrgyzstan. *European Scientific Journal*, 3, 237-245. Retrieved from http://eujournal.org/index.php/esj/article/viewFile/2941/2770

Kushner, D. (2013, February 26). The Real Story of Stuxnet. *IEEE Spectrum, 50*(3). Retrieved from http://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet

Lee, J. K. (2015). Guest editorial: Research framework for AIS grand vision of the bright ICT initiative. *MIS Quarterly*, *39*(2), iii-xii.

Lee, J. K., Cho, D. G., & Lim, G. G. (2018). Design and Validation of the Bright Internet, *Journal of the Association for Information Systems, 19*(2), 63-85

Lee, S. (2016, March 23). Ransomware Wreaking Havoc in American and Canadian Hospitals, *Newsweek.* Retrieved from http://europe.newsweek.com/ransomware-wreaking-havoc-american-and-canadian-hospitals-439714?rm=eu

Lee, Y. J. (2011, August 11). SK Comm. Internal networks were infected [in Korean], Digital daily. Retrieved from http://www.ddaily.co.kr/news/article.html?no=81180

March, S. T., and Smith, G. F. (1995). Design and natural science research on Information technology. *Decision Support Systems* 15.4, 251-266.

Milliken, M. & Oatis, J. (2014, December 9). Cyber attack could cost Sony studio as much as $100 million. *Reuters*. Retrieved from http://www.reuters.com/article/us-sony-cybersecurity-costs-idUSKBN0JN2L020141209

Ministry of Science, ICT & Future Planning. (2013, April 10). Korea Joint Investigation Team, "3/20 Cyber-terror Interim Investigation Result Announcement" [in Korean]. Retrieved from http://msip.go.kr/SYNAP/skin/doc.html?fn=02eb95c11fdd4ae7cca0b36632da50d6&rs=/SYNAP/sn3hcv/result/201706/

Mitra, S., & Ransbotham, S. (2015). The effects of vulnerability disclosure policy on the diffusion

of security attacks, *Information Systems Research,* 26(3), 565-584.

Musil, S. (2014, December 4). Sony hack leaked 47,000 Social Security numbers, celebrity data. *Cnet*. Retrieved from https://www.cnet.com/news/sony-hack-said-to-leak-47000-social-security-numbers-celebrity-data/

Nazario, J. (2009). Politically motivated denial of service attacks. *The Virtual Battlefield: Perspectives on Cyber Warfare*, 163-181.

Network World. (2012, November 5). Coca-Cola hacked by Chinese and kept it a secret. Retrieved from http://www.networkworld.com/article/2223443/microsoft-subnet/coca-cola-hacked-by-chinese-and-kept-it-a-secret.html

Nichols, M. (2014, July 7). North Korea complains to UN about film starring Rogen. *Reuters.* Retrieved from http://www.reuters.com/article/2014/07/09/us-northkorea-un-film-idUSKBN0FE21D20140709

Nygren, E. (2015). The Three years since world IPv6: Strong IPv6 Growth Continues. *The Akamai Blog* [Weblog post]. Retrieved from https://blogs.akamai.com/2015/06/three-years-since-world-ipv6-launch-strong-ipv6-growth-continues.html

Oetzel, M. C., & Spiekermann, S. (2014). A systematic methodology for privacy impact assessments: A design science approach, *European Journal of Information Systems, 23*(2), 126-150.

Oh, B. M. (2011, April 6). Police said, "3.4DDoS attack and 7.7DDoS attack, same attacker." *BoanNews*. Retrieved from http://www.boannews.com/media/view.asp?idx=25581

Park, N. H. & Chung, M. H. (2016). Discussion analysis on the 4th UNGGE about information security and development prospects [in Korean]. *National Strategy*, *22*(3), 173-198.

Park, S. C. (2013, January 25). The whole world's one million computers conducted DDoS attack on Estonia with 1.3 million citizens [in Korean]. *ChosunBiz*. Retrieved from http://biz.chosun.com/site/data/html_dir/2013/01/24/2013012402645.html

Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go? *MIS Quarterly*, *35*(4), 977-988.

Ponemon Institute. (2015). 2015 cost of cyber crime study: Global. Retrieved from http://www.cnmeonline.com/myresources/hpe/docs/HPE_SIEM_Analyst_Report_-_2015_Cost_of_Cyber_Crime_Study_-_Global.pdf

Ponemon Institute. (2017). 2017 cost of cyber crime study. Retrieved from https://www.accenture.com/us-en/insight-cost-of-cybercrime-2017

Porras, D. A. (2006). Comment: The "common heritage" of outer space: Equal benefits for most of mankind, *California Western International Law Journal, 37*(1), 143-176.

Rehman, S. (2013, January 14). Estonia's lessons in cyberwarfare. *U.S. News and World Report*. Retrieved from http://www.usnews.com/opinion/blogs/world-report/2013/01/14/estonia-shows-how-to-build-a-defense-against-cyberwarfare

Roscini, M. (2010). World wide warfare: Jus ad bellum and the use of cyber force. *Max Planck Yearbook of United Nations Law 14,* 87-88.

Ruus, K. (2008). Cyber war I: Estonia attacked from Russia. *European Affairs, 9*(1-2). Retrieved from http://www.europeaninstitute.org/index.php/component/content/article?id=67:cyber-war-i-estonia-attacked-from-russia

Sanger, D. E. (2012, June 1). Obama order sped up wave of cyberattacks against Iran. *The New York Times*. Retrieved from http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html

Sanger, D. E. & Fackler, M. (2015, January 18). N.S.A. Breached North Korean Networks Before Sony Attack, Officials Say. *The New York Times*. Retrieved from https://www.nytimes.com/2015/01/19/world/asia/nsa-tapped-into-north-korean-networks-before-sony-attack-officials-say.html?_r=1

Schmitt, M. N. (ed.). (2017). *The Tallinn Manual 2.0 on the international law applicable to cyber operations*. Cambridge, UK: Cambridge University Press [Kindle Version]. Retrieved from www.amazon.com.

Sherstobitoff, R., & Itai Liba, M. (2013). *Dissecting operation Troy: Cyberespionage in South Korea* (McAfee White Paper). Retrieved from https://paper.seebug.org/papers/APT/APT_CyberCriminal_Campagin/2013/dissecting-operation-troy.pdf

Shin, Y. Y., Jeon, S. H., Lim, C. H., & Kim, M. C. (2013). Economic damages assessment for national cyber security measures: Analysis of the March 20 cyber attack. *Journal of National Intelligence Studies*, *6*(1), 129-173. Retrieved from http://www.kanis.or.kr/sample/board_view.php?bbs_id=magazine_search&kbbs_doc_num=65&page=4&searchtype=&searchword=

Sigholm, J. (2013). Non-state actors in cyberspace operations. *Journal of Military Studies*, *4*(1).

South Korea, Anti-Terrorism Act of 2016, Pub. L. No. 14071. (2016).

Stefanek, A. (2016). Memo from the CEO, Hollywood Presbyterian Medical Center. Retrieved from http://hollywoodpresbyterian.com/default/assets/File/20160217%20Memo%20from%20the%20CEO%20v2.pdf

Steinbart, P. J., Keith, M. J., & Babb, J. (2016). Examining the continuance of secure behavior: A longitudinal field study of mobile device authentication, *Information Systems Research, 27*(2), 219-239.

Supreme Prosecutors' Office (2015). Interim Investigation result on KHNP Cyber-terror Incident [in Korean]. Retrieved from http://www.spo.go.kr/spo/notice/press/press.jsp?mode=view&board_no=2&article_no=593028

Symantec. (2014). Malware discovered in Sony hacking has links to attacks on South Korea [in Korean]). Retrieved from http://www.yonhapnews.co.kr/bulletin/2014/12/08/0200000000AKR20141208109000017.HTML

The FBI. (2014). Update on Sony investigation. Retrieved from https://www.fbi.gov/news/pressrel/press-releases/update-on-sony-investigation

Tikk, E., Kaska, K., & Vihul, L. (2010). *International Cyber Incidents Legal Considerations.* Tallinn, Estonia: Cooperative Cyber Defence Center of Excellence. Retrieved from https://ccdcoe.org/publications/books/legalconsiderations.pdf

United Naitons (UN). (1945). Charter of the United Nations Retrieved from http://www.un.org/en/charter-united-nations/

United Nations (UN). (2015). Report of the group of governmental experts on developments in the field of information and telecommunications in the context of international security, A/70/174. Retrieved from http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174

United Nations Office for Disarmament Affairs. (1968). Treaty of the Non-Proliferation of Nuclear Weapons. (1968). *United Nations Office for Disarmament Affairs*. Retrieved from https://www.un.org/disarmament/wmd/nuclear/npt

United States Cybersecurity Information Sharing Act of 2015. Pub. L. No. 114-113, 129 Stat. 694 (2015).

Wang, J., Xiao, N., & Rao, H. R. (2015). An exploration of risk characteristics of information security threats and related public information search behavior, *Information Systems Research*, *26*(3), 619-633.

Willison, R., & Warkentin, M. (2013). Beyond deterrence: An expanded view of employee computer abuse, *MIS Quarterly, 37*(1), 1-20.

Zetter, K. (2011, July). How digital detectives deciphered Stuxnet, the most menacing malware in history. *Wired*. Retrieved from https://www.wired.com/2011/07/how-digital-detectives-deciphered-stuxnet

Zetter, K. (2016, March). Why hospitals are the perfect targets for ransomware, *Wired*. Retrieved from https://www.wired.com/2016/03/ransomware-why-hospitals-are-the-perfect-target

Appendix

Table 1. Cases of State-Led Cyberattacks

Cases	(S1) Sony hacking	(S2) Iran Stuxnet attack	(S3) Georgia DDoS attack	(S4) Estonia DDoS attack	(S5) KHNP hacking	(S6) 3/20 APT attack	(S7) 7/7 DDoS attack
(1) Purposes of cyberattacks	Cancellation of showing of "The Interview" movie. (a)	Progress delay of Iran nuclear development. (a)	Maintaining the lead before and during war.	Expression of dissatisfaction with the movement of a monument.	Causing social chaos and anxiety.	Causing social chaos and damage. Gaining intelligence regarding the targets to carry out further attacks. (a)	Political attack. (a) Causing disruption, rather than stealing data. (b) Gathering Korea and U.S. military information. (c)
(2) Targets of cyberattacks	Sony Pictures Entertainment (Sony)	Iran Natanz Nuclear facility (German SCADA system)	Homepage of Georgian President, parliament, national defense, foreign ministry, banks, press, portal, Western press and websites of Georgian hackers. (a)	Information distribution channels, national Internet infrastructure, governmental and political websites, commercial & private services, emergency numbers, 112, etc.	KHNP computer system and data	At least 20 organizations including three major banks and three broadcasters.	14 U.S. government institutions and 22 Korean government institutions, etc. (d)
(3) Origin country	Shenyang, China (North Korean hackers known as "Guardians of Peace"). (b)	Prepared by Israel and U.S. (a)	Russia (Russian Business Network)	Russia (Russian government IPs)	North Korea (North Korean IP) Beijing, China (IP of North Korea Posts and Telecom Co. [KPTC]) Shenyang, China (Chinese IP) Russia (Russian IP)	North Korea (North Korean IP)	China (North Korean KPTC' s IP). (e)

Table 1. Cases of State-Led Cyberattacks

(4) Attack means and methods	Malware "Destover." (c) Sony intranet was infected with malware via an e-mail attack and compromised system was used to propagate malicious code. (d) File deletion, master boot record destruction through shared networks sequentially. (d) Hacking message displayed on PC screens of Sony (e).	Malware "Stuxnet." Nuclear facility innersystem compromised through a universal serial bus (USB) flash drive.	DDoS. Traffic blockade through attacks on routers in Georgia. Major banks' systems paralyzed. Defacement of websites. Domain name system (DNS) servers attacked.	DDoS. Critical infrastructure paralyzed for two weeks by DDoS attack. Six instances of blackmail. Master boot record destruction.	Kimsuky affiliation malware. Known to be used by North Korean hackers. E-mail attacks (5,986 phishing e-mails to 3,571 employees).	APT attacks. Attacked four times. Programmed attack time-loaded SQL injection, web shell upload Careful preparation: 76 types of malicious codes, structures and operations were different for each company attacked. (b)	Malware, web hard sites exploit. Attacked four times. Cooperation between malicious codes. Attacks on multiple domains. Damage to important files and documents. Hard disk damaged by software. After attack, zombie PCs conducted self-destruction.
Attack characteristics	Spear phishing attack exploiting zero-day vulnerability. (f)	Typical man-in-the -middle attack.	Structured query language (SQL) injection. Zombie computers used. Russian hackers attacked Internet connection passage to paralyze the Georgian Internet.	SQL Injection with DDoS attack. About 1 million computers were mobilized.	Social engineering through a watering hole using a social media website such as Twitter.	Spread of malicious codes using a vaccine update program. About 48,000 PCs were mobilized.	Hacking and compromising web service to distribute malware that programmed attack time. 115,044 PCs were mobilized.
(5) Attack routes	Direct connection to target. (b)	Infected USB - C&C server.	Through servers located in Russia and Turkey.	Approximately 100 countries. (a)	Direct connection and routed through at least five countries.	49 routers (domestic 25, overseas 24). Using VPN.	Communication with 435 servers of at least 61 nations.
(6) Attack timing	Before showing the film.	Appropriate time for using an infected USB memory stick.	Before and during the war.	Appropriate time for instigating pro-Russian emotion.	Appropriate time for attacking.	Appropriate time for attacking	Independence Day to draw attention.

Table 1. Cases of State-Led Cyberattacks

(7) Attack duration	Early Sept.(b) –Nov. 24, 2014 (g)	Jan. 2010– Sept. 2010 (b)	July 19–Aug. 28, 2008	Apr. 27– May 19, 2007	Dec. 9, 2014– Mar. 12, 2015	Mar. 20–26, 2013	1st Attack: July.4 (EST), 2nd Attack: July 7 (KT), 3rd Attack: July 8 (KT) 4th Attack: July 9 (KT).
(8) Preparation period	About three months (b)	Bush admin. – Dec. 2009	Unidentified	Unidentified	At least six months	At least nine months	Five months (f)
(9) Investigation period	26 days	Unidentified	Unidentified	Unidentified	159 days	21 days	21 days
(10) Consequences	Sony's reputation damaged by breach US$100 million. (h)	Caused nearly two-year delay in nuclear development. (a)	Georgia surrendered after five days of war with Russia.	Tens of millions US$. (b) Administration tasks paralyzed for two months.	Social unrest. Fall of image for nuclear exporting countries.	Social unrest US$867.2 million. (c)	Cybersecurity systems' vulnerability exposed. US$36.3– 54.4 million. (g)
Damage in detail	47,000 cases of personal information leaked. (g) 33,000 cases of internal company documents leaked. (g) Five unopened movies leaked. (g)	Stuxnet may have destroyed about 1,000 centrifuges.	Lost control of national domain names. Banking operations paralyzed.	58 sites' services interrupted for two months.	Eight PCs infected, 94 files leaked, including the nuclear power plant's blueprints and personal information.	48,700 PCs were infected, 32,552 servers and PCs damaged, and data destroyed. 16,221 units of CD/ ATMs damaged and data destroyed. (c)	1,300 units of PCs damaged, 11 million units of PCs infected.
(11) Applicable laws	Related domestic laws such as criminal law and the Mutual Legal Assistance Treaty. When the results of the attack necessitate a use of force, the attack should face the UN Security Council and the international court.						

Table 1. Cases of State-Led Cyberattacks

(12) References	*The FBI, 2014 (a) Nichols, 2014; Gallager, 2014 (b) Sanger & Fackler, 2015 (c) Symantec, 2014 (d) Gil, 2014 (e) Daily Mail, 2014 (f) Hesseldahl, 2015 (g) Musil, 2014 (h) Milliken & Oatis, 2014	* Zetter, 2011 (a) Sanger, 2012 (b) AhnLab, 2010	* Tikk et al., 2010 (a) Kozlowski, 2014	* Tikk et al., 2010 (a) Park, 2013 (b) Kim, 2013	* Supreme Prosecutors' Office, 2015	* Ministry of Science, ICT & Future Planning, 2013 (a) Sherstobitoff & Itai Liba, 2013 (b) Ahnlab, 2013 (c) Shin et al., 2013	* Korean National Police agency, 2009 (a) Nazario, 2009 (b) Choe &John, 2009 (c)Sherstobitoff & Itai Liba, 2013 (d) Chae, 2010 (e) Oh, 2011 (f) Kim, 2015 (g) Emerging Technology Research Center, 2009

Table 2. Cases of Private-Led Cyberattacks

Cases	(P1) Hollywood hospital hacking	(P2) SK Communications hacking	(P3) Coca-Cola hacking
(1) Purposes of cyberattacks	Asking for money: 40 Bitcoins (US$17,000).	Leakage of customer records.	Leakage of negotiation information.
(2) Targets of cyberattacks	Hollywood Presbyterian Medical Center	SK Communications	Coca-Cola Corporation executives' computers
(3) Origin country	Unidentified	China (Chinese hackers)	China (Chinese hackers) (a)
(4) Attack means and methods	Hackers used malware to infect a hospital's computers and to prevent hospital staff from communicating with those devices. Using ransomware called "Locky" to lock systems by encrypting files. (a) An employee mistakenly clicked on an e-mail attachment that was actually a phishing scam. (b)	Turning the company's internal network PCs into zombies, obtaining the manager's authorization, connecting to the database server, and leaking data with a Chinese IP through remote control . After hacking "Alzip" update server of Estsoft, the hackers spread malicious codes .	Phishing e-mails to executives of the Coca-Cola Company. Attackers installed a keystroke logger to steal e-mails, documents and computer account passwords for easy access to the company's network. (b)

Table 2. Cases of Private-Led Cyberattacks

Attack characteristics	Encryption of patient information for money, and the hackers keep the ransom by providing a decryption key. (c)	Hacking for a targeted attack (a precise spread of malicious code aimed at a specific site). (a) 62 Inner infected PCs were mobilized.	Coca-Cola did not know until the FBI informed them that they were hacked. Coca-Cola kept the breach a secret due to concerns over stock prices and credibility. (b)
(5) Attack routes	Unidentified	Indirect attack through the US	Unidentified
(6) Attack timing	Selecting and attacking targets after finding a loophole in security.	Usual time for gathering information.	Before and during negotiations.
(7) Attack duration	Feb. 5– 15, 2016	July 26, 2011	Feb.16– Mar. 16, 2009
(8) Preparation period	Unidentified	At least eight days (July 18– July 25)	Unidentified
(9) Investigation period	Unidentified	15 days	Unidentified
(10) Consequences	Discrediting of hospital name.	Fall of corporate image.	Fall of corporate image.
Damage in detail	US$17,000	Leakage of 35 million cases of customer information .	Leakage of inside information and reputation damaged . Acquisition plan of the China Huiyuan Juice Group fell through.
(11) Applicable laws	Domestic related laws such as criminal law and international law through the Mutual Legal Assistance Treaty.		
(12) References	* Stefanek, 2016 (a) Zetter, 2016 (b) Infosec institute, 2016 (c) Lee, 2016	*Korean National Police Agency, 2011 (a) Lee, 2011	* Charette, 2012 (a) BBC News, 2012 (b) Network World, 2012

Table 3. Origins, Paths, and Targets of SLCAs and PLCAs

Cases	Origins	Paths	Targets	References
(S1) Sony Hacking (2014)	Chinese IPs (exclusively used by North Korean) (a) →	→ Chinese ISP → U.S. ISP →	→ Sony	*The FBI, 2014 (a) Sanger & Fackler, 2015
(S2) Iran Stuxnet Attack (2011)	An infected USB flash drive (Man-in-the Middle Attack) →	→ The malicious worm infiltrated into the inner network →	→ Natanz's nuclear facility equipment was infected by a virus, which was controlled by C&C servers in Denmark and Malaysia → Compromised process-control network of intra-network through OS server → Control system network	*Hruska, 2015 *Zetter, 2011 *Csanyi, 2011
(S3) Georgia DDoS attack (2008)	Russian IP →	→ Russian ISP → Turkish, Ukrainian ISP → Georgian ISP →	→ Governmental organizations' home pages	*Tikk et al., 2010
(S4) Estonia DDoS attack (2007)	Russian IP →	→ Russian ISP → One million zombie PCs were mobilized from approximately 100 countries. (a) → Estonian ISP →	→ Critical infrastructure	*Tikk et al., 2010 (a) Park, 2013
(S5) KHNP hacking (2014–2015)	25 North Korean IPs →	→ North Korean ISP (SJV)(a) → Chinese ISP → South Korean VPN (illegal use of another´s name)→	→ Nuclear power plant (log record was confirmed)	* Korea Supreme Prosecutors' Office, 2015
	Five IPs of North Korean Post and Telco (KPTC) located in Beijing China →	→ Chinese ISP → South Korean VPN (a North Korean IP was found on a South Korean VPN) → South Korean ISP →	→ Nuclear power plant (log record was confirmed)	
	IP(s) from Shenyang, China (175.167.xxx.xxx) →	→ Chinese ISP → ISPs in the USA, China, Japan, Thailand, and The Netherlands → South Korean VPN → South Korean ISP →	→ Collected information from a retiree community → Sent phishing e-mails to retired employees → Collected and leaked information from the affiliated companies → Threatened five times by blackmail to extort money and to reveal the leaked information to the public through Naver, Twitter, and other routes.	

Table 3. Origins, Paths, and Targets of SLCAs and PLCAs

	Russian IP (Vladivostok) →	→ Russian ISP → South Korean ISP →	→ Leaked the blueprints of the nuclear power plant and exposed through a Twitter account that was used by a North Korean hacker, and attempted extortion.	
(S6) 3/20 APT attack (2013)	North Korean IP →	→ North Korean ISP → Chinese ISP (VPN) → South Korean ISP → PMS Server → Zombie PCs → Overseas ISPs → C&C Servers in four detour states in the first attack and seven detour states in the second attack → Detour state's ISPs→ South Korean ISP →	→ Financial institutions, automated teller machines, press, and servers in South Korea	* Ministry of Science, ICT & Future Planning, 2013
(S7) 7/7 DDoS attack (2009)	North Korean KPTC's IP in China →	→ Chinese ISP → South Korean ISP → South Korean file sharing sites → Contamination of malware and production of zombie PCs in 61 countries and 435 servers → Detour states' ISPs →	→ 14 organizations in the US → 22 organizations in Korea including governmental agencies, the press, and civilian enterprises	* Clarke & Knake, 2010 * Korean National Police agency, 2009
(P2) SK communications	Chinese IP	→ Chinese ISP → U.S. ISP → U.S. Domain → U.S. ISP→ Korean ISP → hacking Antivirus update server in Korea	→Infection of 62 PCs of SK Communications (becoming Zombie PCs)	* Korean National Police Agency, 2011
* The attack routes of the (P1) Hollywood Hospital and (P3) Coca-Cola cases were not identified.				

About the Authors

Young Yung Shin is a senior researcher at the Bright Internet Research Center at the Korea Advanced Institute of Science and Technology (KAIST) and a vice president of the Korea National Cyber Security Association. He received an MA in international relations from the Fletcher School at Tufts University in 2012, and a PhD in computer engineering (information security) from KAIST in 2017. His research interests include the Internet Peace Principles, cyberlaws, institutions and policy, national cybersecurity, and global Internet governance. He has published journal articles and presented conference papers in the area of cybersecurity. He currently serves as an auditor general of the Korea Transportation Safety Authority (KTSA). Prior to joining KTSA, he was a director general in the Presidential Security Service of the Republic of Korea for 28 years.

Jae Kyu Lee has been a professor and chair professor at KAIST since 1985, and since September 2016 a professor emeritus of KAIST. During this research, he was the Wang Yingluo Fellow Professor at Xi'an Jiaotong University and a distinguished visiting professor at Heinz College at the Carnegie Mellon University in Pittsburgh. He is a fellow and past president of the Association for Information Systems and is founder of the Bright Internet and the Bright Internet Global Summit. He was a conference cochair of many international conferences including the International Conference of Information Systems 2017. He received his MS from KAIST, and his PhD from the Wharton School at the University of Pennsylvania. He recently joined the Southern University of Science and Technology in Shenzhen as visiting chair professor.

Myungchul Kim received a PhD in Computer Science from the University of British Columbia, in Vancouver, Canada, in 1993. He has been a professor at KAIST since 2009. From 1984-1997 he was a managing director in the Korea Telecom Research and Development Group, where he was in charge of research and development of protocol and QoS testing on ATM/B-ISDN, IN, PCS, and the Internet. He has also served as a program committee member at many conferences, including IWTCS, IEEE ICDCS, and IFIP FORTE, and was cochair of the IWTCS '97 and the FORTE '01. He has published over 150 conference proceedings, book chapters, and journal articles in the areas of computer networks, wireless mobile networks, protocol engineering, and network security.