MAY 2026

Building a Robust U.S.-ROK Cyber Alliance

A Joint Cyber Resilience Strategy

PROJECT DIRECTORS

Donghee Kim

James Andrew Lewis

AUTHORS

Sunha Bae

Julia Brock

Donghee Kim

Yena Kim

James Andrew Lewis

Joohui Park

A Report of the CSIS Strategic Technologies Program

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES | NSR National Security Research Institute

MAY 2026

Building a Robust U.S.-ROK Cyber Alliance

A Joint Cyber Resilience Strategy

PROJECT DIRECTORS

Donghee Kim

James Andrew Lewis

AUTHORS

Sunha Bae

Julia Brock

Donghee Kim

Yena Kim

James Andrew Lewis

Joohui Park

A Report of the CSIS Strategic Technologies Program

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

Center for Strategic & International Studies
1616 Rhode Island Avenue, NW
Washington, DC 20036
202-887-0200 | www.csis.org

Cover: Ruma via AdobeStock

Acknowledgments

The authors of this report would like to thank the experts and practitioners who provided key insights for this research through roundtables and events. The authors would also like to thank Kuhu Badgi for her excellent management support for this project. Finally, the authors are thankful to the CSIS publications team for their help with report editing and publication.

This report is made possible by general funding to CSIS and generous support from South Korea's National Security Research Institute.

Contents

Executive Summary

South Korea (ROK) is at the forefront of global cyber conflict, facing intensifying threats from North Korea (the DPRK), China, and Russia. In this dynamic threat environment, cybersecurity has become a critical pillar of South Korea's national security and a key element of the U.S.-ROK alliance. Recognizing the challenge, South Korea and the United States signed the Strategic Cybersecurity Cooperation Framework (SCCF) in 2023, reaffirming and expanding bilateral cooperation in areas such as cyber defense, intelligence sharing, critical infrastructure protection, and active countermeasures against malicious cyber actors. The SCCF is expected to be reviewed and updated later this year, and translating this high-level commitment into concrete operational frameworks and coordinated action remains an ongoing challenge.

This joint research project between the Center for Strategic and International Studies and the National Security Research Institute of Korea establishes foundational frameworks to strengthen bilateral cyber defense and resilience.

The report highlights several key areas for action:

- **Strengthening shared situational awareness essential.** This includes improving threat identification, risk assessment, and timely, actionable information-sharing mechanisms between South Korea and the United States, as well as the development of the Cyberattack Severity Classification Framework.
- **Attribution must be understood as both technical and policy driven.** In the case of North Korea, where cyber operations are strategically directed by the state, attribution

requires close coordination between two countries as well as stronger joint capabilities to enable timely, credible, and coordinated responses.

- **South Korea needs to adopt more proactive approaches.** This report identifies the limitations of existing defensive and reactive approaches and emphasizes the need to shift toward more proactive strategies. Rather than remaining at the level of general principles, it outlines the response options and institutional foundations needed for South Korea to operationalize such approaches in practice.
- **North Korea's cyber activities require a more integrated response.** North Korea's cyber activities are becoming increasingly complex and transnational, extending beyond the cyber domain into financial systems, legal jurisdictions, and global labor markets. As a result, purely technical or cyber-centric responses are insufficient. Addressing these challenges requires integrating cyber capabilities with law enforcement, financial regulation, and international cooperation, including the use of legal frameworks such as the UN Convention against Cybercrime.

Building on this analysis, the report emphasizes that current approaches remain insufficiently integrated to address transnational threats in practice. Addressing these challenges will require expanded collaboration with additional allies, partners, and private-sector stakeholders. In particular, this requires more structured cooperation frameworks and clearer divisions of roles among stakeholders.

Introduction

The cybersecurity environment on the Korean Peninsula has become a major security challenge. Being neighbors with Russia, China, and North Korea makes cyber threats a constant risk. In particular, North Korea uses cyber activities as a source of state revenue, enabling sustained pressure beyond conventional military tensions.

These threats have undergone qualitative changes in recent years. Cyberattacks can now take the form of continuous activities combining financial theft, espionage, and influence operations. In the case of North Korea, cross-border networks have been expanded through cryptocurrency theft and the use of foreign-based IT workers. The transnational nature of these activities complicates attribution and response while enabling the exploitation of gaps in international cooperation. These activities simultaneously pursue revenue generation and sanctions evasion. In this process, the boundary between cybercrime and state behavior is deliberately blurred.

The international environment also shows clear limitations. The 2025 UN Open-Ended Working Group (OEWG) final report reaffirmed existing norms but did not bridge the gap between norms and operational implementation. As discussions continue under the UN Global Mechanism on Cybersecurity, meaningful progress in developing international law and norms is expected to remain limited.

These conditions highlight the limitations of traditional defensive approaches in cyberspace. Previous strategies have focused on defensive measures and reactive responses, but these efforts have failed to impose meaningful constraints on adversaries. In addition, cyber activities have

often been treated as a separate domain, without sufficient integration into the broader context of interstate competition and conflict.

Many countries have attempted to develop cyber deterrence policies, but these have failed to establish credible threats or achieve effective deterrence. Heightened concerns about escalation and exaggerated demands for attribution accuracy have often complicated decisionmaking and hindered timely and effective responses. As a result, cyber operations are still widely perceived by attackers as low cost and low risk. This is particularly pronounced in the case of North Korea. While deterrence functions in the conventional military domain, deterrence imposes few effective limits on cyber activities in the gray zone below the use-of-force threshold.

North Korea is unlikely to conduct cyber operations above the use-of-force threshold, and it could be argued that its current cyber activities should be treated as crime or political coercion and tolerated to some extent. However, this view is becoming increasingly inconsistent with national security. As dependence on digital infrastructure in South Korea grows, the impact of cyberattacks expands. In addition, given North Korea's decisionmaking structure, it is difficult to rule out the possibility that limited activities could escalate rapidly. Unchecked cyber activity by North Korea also carries a growing risk of unintended consequences.

U.S.-ROK cooperation has advanced in recent years. At the April 2023 U.S.-ROK summit, the two countries jointly adopted the Strategic Cybersecurity Cooperation Framework (SCCF), expanding their traditional land, sea, and air alliance into cyberspace. Since then, cooperation has expanded through the U.S.-ROK Cyber Policy Consultation to include the protection of critical and defense infrastructure, joint cybercrime investigations, and cyber defense exercises. However, these efforts have largely remained declaratory, and further steps are needed to ensure effective implementation. The framework is expected to undergo a formal "refresh" later this year or in 2027.

These issues ultimately lead to a more fundamental question of how South Korea can build a strong cyber defense posture and capabilities, as well as generate credible threats in cyberspace. This requires a comprehensive approach that integrates policy tools and cooperation with allies and partners.

This study seeks to advance a more concrete discussion of these questions. Its starting point is the development of a shared framework for cyber situational awareness. Without common standards for defining threats, assessing the situation, and determining levels of severity, effective cooperation remains difficult.

Attribution remains a critical challenge. It is a technical issue, but also an issue that requires policy judgment. In the case of North Korea, cyber operations are unlikely to occur without state involvement, underscoring the importance of informed policy decisions. Strengthening joint attribution capabilities between the ROK and the United States, along with appropriate policy judgment, is a key element of a coordinated cyber response.

On this basis, a shift toward more proactive approaches is required. Active cyber defense involves identifying threats in advance and continuously disrupting adversary activities to

constrain operations and increase costs. This approach requires moving beyond a reactive and defense-centric model toward more persistent and structured engagement. In the case of North Korea, where cyber activities are closely linked to financial gain and regime stability, reducing the effectiveness and increasing the cost of these activities becomes particularly important.

Such responses cannot remain confined to cyberspace. Activities such as cryptocurrency theft and the use of overseas IT workers rely on cross-border networks, requiring an integrated approach that combines law enforcement, financial regulation, and diplomatic measures, as well as close coordination with allies and partners. More broadly, responding effectively also requires integration with other domains as part of a comprehensive strategy to address North Korea.

Ultimately, cyber resilience is not solely a matter of individual capability, but of how effectively systems are connected and coordinated. U.S.-ROK cooperation provides a critical foundation in this regard and should evolve toward more integrated and coordinated defense and response while expanding cooperation with additional allies and partners.

This report builds on these considerations and presents policy recommendations for strengthening South Korea's cyber resilience, with a focus on U.S.-ROK cooperation. It outlines key areas for improving shared situational awareness, advancing South Korea's active cyber defense, and identifying available response options.

Part I

Common Criteria for U.S.-ROK Cyber Situational Awareness

Criteria for Cyber Situational Awareness

By Julia Brock and James Andrew Lewis

Cyber threats against South Korea and the United States have increased in recent years, prompting both nations to strengthen cooperation on cyber issues. South Korean domestic public institutions saw a 36 percent **increase** between 2022 and 2023 to 1.6 million cyberattacks, and the number of detected cyberattacks against U.S. targets **surged** 136 percent between October 2024 and April 2025. Both countries have enhanced their cyber cooperation to combat these cyber risks through the **U.S.-ROK Joint Leaders' Statement in 2022** and the **declaration** that the U.S.-ROK mutual defense treaty extends to cyberspace in 2024. Both countries must establish criteria for cyber situational awareness to maintain robust cyber cooperation. Cyber situational awareness aims to provide a comprehensive view of the environment, threats, and malicious activity that can form the basis for action in cyberspace. Maintaining **awareness** of this environment requires continuous monitoring, analysis, and evaluation of cyber activities to identify vulnerabilities, threats, and the potential effects on the economy, society, and national security.

A cooperative approach to cyber situational awareness between the United States and South Korea will require both nations to develop processes and agreed standards for coordinating real-time communication, to make more timely and better-informed decisions, and to take proactive measures to defend digital assets and infrastructure from cyber threats. The categories of information needed for cyber situational awareness include:

- threat intelligence;
- network monitoring;
- vulnerability assessments;
- incident detection and analysis;
- all-source risk assessment; and
- near real-time situational reporting.

At its core, cyber situational awareness involves three fundamental dimensions: (1) acquiring knowledge about what is happening in the digital environment, (2) understanding and being able to explain why these events are occurring, and (3) assessing what impacts these developments could have on national security, economic stability, and public safety. This process necessitates sophisticated data collection methodologies, advanced analytical frameworks, and well-organized information management systems. Most importantly, to establish cyber situational awareness, there must be agreement between both nations regarding information sharing protocols, task allocation procedures, and prioritization frameworks, using the existing framework of collaborative efforts by the two countries.

Data Collection and Intelligence Infrastructure

Establishing effective cyber situational awareness first requires producing capabilities for collecting and sharing aggregate data from diverse sources. This data should cover government networks, critical infrastructure sectors (e.g., energy, finance, healthcare, transportation, and water systems), private companies, academic institutions, and open-source intelligence. The data must be processed to identify malicious activities and understand the tactics and techniques of various threat actors, including state-sponsored entities, sophisticated cybercriminal organizations, and ideologically motivated hacktivists.

AI Tools

Artificial intelligence (AI) tools can improve cyber situational awareness by processing and analyzing vast amounts of security-related data at speeds and scales beyond human capabilities. This enables organizations to gain a more comprehensive understanding of their security posture and the evolving threat landscape. The United States and South Korea can jointly develop algorithms to identify patterns and anomalies in network traffic and (where appropriate) user behavior that might indicate malicious activity with greater accuracy and speed than traditional rule-based systems. AI can also automate parts of the incident response process and enhance vulnerability scanning by prioritizing vulnerabilities. The effective implementation of AI for cyber situational awareness requires a combination of specialized tools supported by skilled security analysts who can interpret AI-produced insights for appropriate action.

Proactive Defense and Predictive Capabilities

Beyond simply reacting to immediate threats, comprehensive cyber situational awareness involves analyzing trends and patterns to predict future cyberattacks and emerging threats, and should be done collaboratively between the United States and South Korea. This predictive capability is essential for proactive cyber defense as it allows for preemptive defensive measures rather than purely reactive responses. The foundation of effective situational awareness between these allied nations relies on robust mechanisms for sharing cyber threat intelligence and information in a timely, secure, and actionable manner.

Shared Analytical Standards

In creating a common approach to situational awareness, the United States and South Korea would benefit substantially from shared analytical criteria designed to identify patterns, relationships, and potential incidents across both nations' digital ecosystems. These analytical frameworks could be integrated with other threat intelligence sources to create a more comprehensive understanding of the **threat landscape** facing South Korea. The information generated should include both high-level strategic summaries for senior decisionmakers and more granular, technically detailed reporting for operational analysts. To facilitate this level of cooperation, South Korea and the United States will need to significantly expand existing communication channels and develop standardized processes for sharing sensitive cybersecurity information. This could build on greater collaboration with the existing National Cyber Awareness System (NCAS) at the Cybersecurity and Information Security Agency (CISA). This tool provides situational awareness to technical and nontechnical audiences by dispensing timely information about cybersecurity threats and issues as well as general security topics. NCAS products include technical alerts, control systems advisories and reports, weekly vulnerability bulletins, and tips on cyber hygiene best practices. The U.S. and ROK governments could consider establishing a similar system between their governments to share data.

Any agreed-upon standards for collective work should establish clear guidelines for timeliness, completeness, and accuracy of shared information. In many cases, real-time or near-real-time data sharing is critical for responding effectively to rapidly evolving cyber threats. All shared data should be consistently formatted and structured to enable effective cross-agency analysis and joint defensive operations. A collaborative effort must include common standards regarding the appropriate level of detail to be exchanged, balancing the need for actionable intelligence with legitimate concerns about sources and methods protection.

These standards and protocols represent critical topics to be jointly developed and agreed upon by cybersecurity authorities from both nations. The creation of joint U.S.-ROK standards for collective cybersecurity must address timeliness to enable rapid defensive actions. These standards should define specific timeframes for sharing different categories of threat intelligence, including immediate notification (within minutes) for critical zero-day vulnerabilities and active attacks against critical infrastructure, and 24-hour windows for less urgent but still significant threat information. The framework should include protocols that automatically accelerate information sharing during crisis situations, such as widespread attacks. Additionally, both nations must

further invest in compatible secure communication channels and automated sharing systems that can transmit encrypted data in standardized formats with minimal human intervention, thereby reducing the delay between threat detection and defensive response.

Joint standards should establish minimum data requirements for different types of cyber threat intelligence to give a full picture of the shared information. These requirements must specify which technical indicators (e.g., IP addresses, malware signatures, and command-and-control infrastructure) and contextual information (targeting patterns, adversary techniques, and potential impacts) must be included in various categories of threat reports. The standards should also clarify expectations about the inclusion of raw data versus analytical conclusions, as well as delineate circumstances under which certain details may be withheld due to classification concerns or source protection. To ensure proper implementation, both countries should create joint review mechanisms to periodically assess the completeness of shared intelligence and identify systematic gaps or areas for improvement in information exchange practices. Both countries will also need to develop common approaches to privacy and personally identifiable information.

With respect to accuracy in reporting, joint U.S.-ROK standards must develop and implement rigorous verification protocols for shared cyber threat intelligence. These protocols could include confidence ratings for different types of information. They should also involve clear sourcing requirements (using a common format for referencing) and procedures for differentiating between confirmed facts and analytical judgments derived from those facts. Both nations should establish joint technical working groups to validate significant technical findings before they trigger major defensive actions while maintaining the ability to rapidly share time-sensitive intelligence with appropriate caveats. Additionally, the standards should include feedback mechanisms that allow recipients to report on the actionability and accuracy of received intelligence, creating a continuous improvement cycle. This focus on accuracy must be balanced with timeliness requirements through established procedures for sharing preliminary information with clear uncertainty markers, followed by more thoroughly verified updates as additional confirmation becomes available.

Strategic Importance

In today's digitally interconnected domain, the United States and South Korea effectively share a border in cyberspace, making improved intelligence sharing not only beneficial but essential for mutual defense. By establishing a comprehensive framework for cyber situational awareness, both nations can enhance their collective ability to detect, analyze, and respond to sophisticated cyber threats targeting their shared strategic interests and critical infrastructure systems.

Beyond countering specific threats, U.S.-ROK cyber cooperation involves extensive information and best practices sharing, facilitated by agreements like the **memorandum of understanding** between CISA and South Korea's National Intelligence Service. This includes collaboration on cyber crisis management, critical infrastructure resilience, and policies related to emerging technologies. While there are challenges such as differing threat perceptions and difficulties in operationalizing active cyber defense on a bilateral basis, future U.S.-ROK cybersecurity collaboration should aim for deeper integration through proactive responses.

Julia Brock *is a former program manager and research associate with the Strategic Technologies Program at the Center for Strategic and International Studies (CSIS) in Washington, D.C.* **James *Andrew Lewis*** *is senior adviser (non-resident) with the Economic Security and Technology Department at CSIS.*

A Cyberattack Severity Classification Framework for the Republic of Korea

By Sunha Bae

Introduction

Despite deterrence efforts, cyber threats continue to escalate, highlighting the need for greater accountability from and cost imposition on malicious actors. The U.S. **2023 National Cybersecurity Strategy** emphasizes imposing costs on malicious actors and reinforcing alliances; the Republic of Korea's (ROK) **2024 National Cybersecurity Strategy** also prioritizes offensive cyber defense and global cooperation. Since 2018, U.S. Cyber Command's "Defend Forward" policy has resulted in 40 **Hunt Forward operations** across 21 countries, exposing threats from major **adversaries**. Similarly, the European Union's updated 2023 **Cyber Diplomacy Toolbox** stresses situational awareness and the importance of holding persistent threat actors accountable.

Attributing cyberattacks and formulating response strategies are inherently political processes shaped by national security priorities, diplomatic relations, and geopolitical **considerations**. Governments must balance deterrence with escalation risks, ensuring proportionality and international legitimacy. It is therefore difficult to establish a single, uniform standard for response. Nevertheless, consistent policy is necessary, as the absence of clear frameworks increases political burdens, delays decisionmaking, and results in inconsistent responses that can confuse allies.

A national framework for classifying cyberattack severity enhances objectivity, guiding policy decisions and facilitating mutual understanding between nations. Although South Korea has shown strong political will to respond to malicious cyber activities, it lacks a clear legal and policy framework for response procedures. To fill this gap, this paper proposes a Cyberattack Severity Classification Framework (CSCF) to objectively assess and categorize cyberattacks, supporting informed decisionmaking.

Purposes of the Cyberattack Severity Classification Framework

- **Support Political Decisionmaking and Strengthen the Framework for Cost Imposition and Accountability:** The CSCF supports the ROK government's political decisionmaking regarding cost imposition and accountability for malicious actors and their state sponsors at the international level. In doing so, the framework helps facilitate sustained international engagement in cyber threat mitigation.
- **Establish a Basis for International Cooperation:** Publicly sharing a national cyberattack severity assessment framework promotes transparency. By aligning with common evaluation criteria, the CSCF supports mutual understanding, helps partners better interpret intent, and strengthens situational awareness, playing a crucial role in advancing international cooperation efforts.

This study, part of a project to enhance ROK-U.S. cybersecurity cooperation, aims to classify cyberattack severity at the national level to support shared understanding and coordinated responses. It reviews government-led models and simplifies criteria around core items for clarity and efficiency. Case studies of state-sponsored cyberattacks in the United States and ROK over the past decade will test the framework's applicability. Although the CSCF is limited by its reliance on public data, it seeks to support cost imposition efforts and to provide a foundation for international cooperation against cyber threats.

National and International Methodologies

UNITED STATES

Presidential Policy Directive 41 (PPD-41): U.S. Cyber Incident Coordination

PPD-41 (2016) aims to enhance national security by managing and coordinating responses to significant cyber incidents. The United States defines a significant cyber incident as a group of related events, rather than a single isolated incident. PPD-41 introduces the Cyber Incident Severity Schema (CISS) to establish a unified federal understanding of incident severity, supporting coordinated national-level responses to significant incidents. CISS outlines six levels—Emergency, Severe, High, Medium, Low, and Baseline—based on the scope of the incidents and their impact on national security and public safety.

Assessment Criteria

- Scope and Impact: Evaluates the level of observed or potential effects on national security, public health, critical infrastructure, the economy, and public confidence
- Observed Actions: Categorizes incidents into preparation, engagement, presence, and effect phases
- Intended Consequence: Assesses attacker intent and potential damage scope

National Cyber Incident Scoring System (NCISS)

Developed by the Cybersecurity and Infrastructure Security Agency (CISA), the **NCISS** provides a consistent method of assessing cyber incidents and prioritizing the response to them across federal agencies. It assigns **risk scores** to incidents, which then guide response priorities and resource allocation. However, the NCISS struggles to assess simultaneous related incidents, relying on expert judgment instead of predefined criteria, which can cause inconsistencies and reduce objectivity in assessments. The NCISS uses the same six-level severity scale outlined by the CISS.

Assessment Criteria

- Functional Impact: Effect on organizational operations
- Observed Activity: Behavior of the threat actor
- Location of Observed Activity: Network areas where malicious activities are detected
- Actor Characteristics: Capabilities and intentions of the threat actor
- Information Impact: Sensitivity of compromised data
- Recoverability: Resources required for incident recovery
- Cross-Sector Dependency: Connections between critical infrastructure sectors
- Potential Impact: Scale and significance of the affected organization

EUROPEAN UNION

NIS Directive 2.0

The NIS Directive 2.0 (2022) establishes a unified framework for network and information security across the European Union. It covers a wide range of critical infrastructure and essential services vital to public safety, economic stability, and national security. Entities are required to assess and report cyber incidents that cause or could cause severe operational disruptions, financial losses, or significant harm to individuals or **organizations**. The scope of critical infrastructure and essential services includes sectors such as energy (electricity, oil, gas, and hydrogen), transport (air, rail, water, and road), banking and financial systems, healthcare, water supply, digital infrastructure (e.g., domain name systems, content delivery networks, and cloud), information and communication technology service management, public administration, and space-related organizations.

Assessment Criteria

- Impact on service operations

- Financial impact
- Impact on other individuals or organizations (material or nonmaterial)
- Importance of affected networks and information systems to the entity's service delivery
- Severity and technical characteristics of the cyber threat
- Extent of impact on service functionality
- Duration of the incident
- Number of affected service recipients
- Whether the entity has experienced similar incidents previously
- Whether the incident was due to malicious or unlawful acts
- Possibility of cross-border impacts

Cybersecurity Incident Taxonomy (CIT)

The **EU CIT**, proposed by the NIS Cooperation Group in 2018, offers a standardized framework for classifying cybersecurity incidents at strategic and political levels. It aims to enhance coordination of incident response activities across the European Union and facilitate cross-border information sharing and cooperation. According to a 2022 Organization for Security and Co-operation in Europe (OSCE) **report**, many European Computer Security Incident Response Teams (CSIRTs) or relevant bodies are utilizing the EU CIT in their processes. Notably, Estonia has adopted the EU CIT as a national model.

Assessment Criteria

- Attack Target: Sector affected by the cyberattack (e.g., energy, transport, finance, healthcare, water, digital infrastructure, government services)
- Threat Severity: Technical risk level of the threat
- Impact Severity: Societal or economic damage level

Cyber Diplomacy Toolbox (CDT)

The **CDT** is an EU framework designed to prevent, deter, and respond to malicious cyber activities, promoting solidarity and mutual support among member states. The CDT emphasizes the importance of shared situational awareness and information sharing among EU member states, aiming to improve the understanding of the European Union's threat environment and encourage the use of the CDT to support decisionmaking. It strengthens coordinated strategies to counter persistent cyber threat actors, reinforces the international obligation of due diligence, and introduces additional response options.

Assessment Criteria

- Scope
- Scale
- Duration

- Intensity
- Complexity
- Sophistication
- Impact

UNITED KINGDOM

Categorization Model for Cyber Incidents (CMCI)

The United Kingdom's National Cyber Security Centre (NCSC) in 2023 used the **CMCI** to classify cyber incidents based on severity and potential impact. It ensures effective resource allocation and response coordination across sectors, including government, critical infrastructure, businesses, and individuals. The NCSC and National Crime Agency (NCA) oversee assessments. Severity levels are classified into six tiers, similar to the U.S. CISS model, based on the scope of affected targets (e.g., nation, institution, individual) and the level of socioeconomic impact.

Assessment Criteria

- National Security Impact: Disruptions to essential services, potential harm to life or national security
- Scale of Disruption: Level of interference in services or operations
- Economic and Social Impact: Financial losses, data breaches, privacy violations, and social consequences
- Attack Target: Categorizes affected entities such as critical infrastructure, government, enterprises, small- and medium-sized enterprises, and individuals

Cyber Regulations 2020 (CR)

The **Cyber Regulations 2020** were established under the Sanctions and Anti-Money Laundering Act 2018 to counter cyber activities threatening the United Kingdom's integrity, economy, and security. These regulations allow the United Kingdom to impose sanctions on individuals or entities involved in cyber threats, including asset freezes and travel bans. Although the regulation does not explicitly define assessment criteria, it outlines the definition of relevant cyber activity and types of cyber activities, which form the basis of the assessment criteria.

Assessment Criteria

- Intent: Whether the attack aimed to cause harm
- Scope of Impact: Number of affected people or systems
- Target: Whether the attack was on critical infrastructure or essential services

FRANCE

National Cyber Attack Classification Scheme (NCACS)

France's Cyber Review (2018) introduced a cyberattack classification system to enhance threat assessment and response coordination. The **NCACS** highlights the need for a clear understanding of

cyberattacks and rapid analysis of attack techniques for proportionate responses. Its classification scheme aims to enhance shared understanding, support decisionmaking, and promote international cooperation on cross-border incidents. The framework prioritizes impact-based evaluations, covering completed or imminent attacks that require urgent responses. Severity levels, compatible with the U.S. CISS, are classified into six tiers based on the scope and impact of cyber incidents. However, the NCACS incorporates international legal considerations, including the possibility of an armed attack under UN Charter Article 51, with the most severe level designated as "possibly assessable."

Assessment Criteria

- Impact: Effect on national interests, security, the economy, and the environment
- Technical Capability: Attack sophistication and methods used
- Intent: Motivations behind the attack
- Target Risk: The affected entity's significance
- Scale: Attack severity and widespread effects
- Repetition: Frequency and relation to past incidents

CANADA

Government Cyber Security Event Management Plan (GC CSEMP)

Canada's Federal Cyber Incident Response Plan differentiates between general and critical incidents, with critical incidents addressed under the **GC CSEMP** (2023). The GC CSEMP ensures coordinated cyber incident management across government systems, led by the Canadian Centre for Cyber Security and the Treasury Board of Canada Secretariat's Office of the Chief Information Officer (TBS-OCIO). It includes the Injury Test, which assesses cyber incident severity and scope, and the Risk Assessment, which evaluates potential exposure. This report focuses on the **Injury Test**, which sorts incidents into four levels based on their scope and severity. Each level is defined by the extent of damage across five areas—public health, financial loss, government services, national security, and reputation—to provide clearer understanding of the attack's severity.

Assessment Criteria

- Severity: Degree of harm across domains such as public health, financial loss, government services, national sovereignty, national security, and reputation
- Scope: Extent of impact in terms of the number of affected individuals, organizations, facilities, or systems; the geographic area (e.g., international, national, multiple sectors, single sector or jurisdiction, individual or small business); and the duration

Australia

Cyber Incident Categorization Matrix (CICM)

The Australian Signals Directorate (ASD) developed the **CICM** to prioritize incident response and ensure appropriate measures. Continuously updated, it is used by the Australian Cyber Security

Centre to classify and manage cyber incidents effectively, as well as provide annual cyber incident **statistics**. Severity levels are classified into six tiers based on the intensity of cyber effects and the significance of the affected organization.

Assessment Criteria

- Cyber Effect: Evaluates impact based on attack success, persistence, and intent (e.g., disruption to critical systems, broad or localized compromise, coordinated or isolated low-level attacks, or failed attacks)
- Significance: Assesses the importance of affected organizations (e.g., the general public, small to large organizations, educational and research institutions, local to federal governments, critical infrastructure, and national security systems)

CHINA

Cybersecurity Incident Classification Guide

To standardize incident reporting and enhance security, China issued the **Cybersecurity Incident Reporting Administrative Measures Draft** (2023) under the Cybersecurity Law, defining **severity levels** and standardizing response procedures for government agencies, critical infrastructure operators, and network service providers. Similarly to GC CSEMP, severity is classified into four levels whose severity is defined across five dimensions—system disruption, data leakage, social impact, economic loss, and harmful information spread—to support rapid decisionmaking.

Assessment Criteria

- System Loss and Functional Disruption: Network failures affecting critical operations
- Information Leakage and Data Loss: Exposure of state secrets and sensitive information
- Social Impact: Number of affected people and essential service users
- Economic Loss: Direct financial damage
- Harmful Information Spread: Dissemination of illegal or harmful content

Summary and Comparison

KEY CHARACTERISTICS OF NATIONAL FRAMEWORKS

The United States assesses cyber incident severity through the CISS framework, integrating PPD-41 and the NCISS. The NCISS uses a quantitative, weighted scoring system to promote consistent assessment practices across the federal government and critical infrastructure sectors. CISA conducts evaluations based on reports from affected entities to ensure consistency.

The European Union classifies incidents via the NIS Directive, CIT, and CDT, sharing core criteria. The CIT, adopted by some member states like Estonia, supports standardized classification and information sharing. Both the NIS Directive and CIT classify attacks on critical infrastructure and essential services as critical threats, clearly identifying relevant sectors

and applying broad coverage. One limitation is that these guidelines—particularly the CIT and CDT—are non-binding recommendations, and therefore not applied consistently across all EU member states.

The United Kingdom classifies cyber incidents through the CMCI and Cyber Regulation, which evaluate attack targets as well as economic, operational, and social impact. The sanctions regulation also considers attacker intent (malicious or unlawful) to ensure accountability and cost imposition on threat actors. Similar to the U.S. system, the NCSC and NCA carry out categorization to promote consistent assessment.

France employs cyber incident severity classification to support proportionate responses to malicious activities, emphasizing situational awareness and international cooperation. The NCACS is designed to be compatible with the U.S. CISS, aiming to enhance mutual understanding of cyber threat levels in the context of international coordination.

Canada's GC CSEMP assesses threats and risks using a matrix-based system to determine response levels. Similarly to the United Kingdom, Canada classifies attack targets by scale and considers the transnational scope of incidents.

Australia's CICM prioritizes victim importance and impact, using a simplified approach. The government maintains and updates its matrix, publishing annual statistics on the severity of cyber incidents in Australia to enhance awareness. The 2023 update categorizes impact by scope, focusing on supply chains, shared services, and critical infrastructure. Australia's framework remains adaptive to changing cyber threats.

China has introduced a cyber incident classification guide and is developing supporting legislation. While it is similar to international models in most assessment criteria, China's system emphasizes controlling harmful information, in line with state information policies. The framework includes quantitative standards, enabling rapid identification and reporting of major cyber incidents, though the basis for these standards is not explicitly provided.

CORE CRITERIA

Across countries, cyber incident assessment criteria generally outline three tiers of importance. Core criteria—such as the scope and importance of the target; attacker intent; impacts on operations, the economy, society, and national security; potential loss of life; and scale and duration—are prioritized in most frameworks. Frequently considered but less emphasized are secondary factors like technical capabilities and incident frequency. Finally, other factors—such as infrastructure dependencies, attack success, harmful information dissemination, and recoverability—are selectively used depending on the context and purpose of the assessment.

While national goals vary, many countries acknowledge the value of severity classification frameworks for crisis response, resource allocation, and international cooperation. Notably, models like the European Union's CDT, the United Kingdom's regulatory framework, and France's NCACS show how classification can support accountability and diplomatic engagement.

Table 1: Comparison of Assessment Criteria

Category	Criteria	Countries/Regions
Attack Target	Scope, importance of system/target	United States (PPD-41, NCISS), European Union (NIS, CIT, CDT), United Kingdom (CMCI), France, Canada, Australia, China
	Cross-sector dependency	United States (NCISS)
Attacker Intent/ Capability	Attacker intent (malicious/illegal)	United States (PDD-41, NCISS), European Union (NIS), United Kingdom (CR), France, Canada, Australia, China
	Complexity/sophistication	United States (PDD-41, NCISS), European Union (NIS, CDT), France
Impact and Damage	Economic, functional, informational, social, and national security impact; risks to health and safety	United States, European Union, United Kingdom, France, Canada, Australia, China—all of which include multiple dimensions, with slight variations
	Scale, duration, intensity	
	Frequency	France
	Recoverability	United States (NCISS)
	Spread of harmful information	China

Source: Authors' analysis.

Cyberattack Severity Classification Framework (CSCF)

ASSESSMENT CRITERIA

The CSCF criteria are based on international frameworks and focus on core criteria to improve efficiency and consistency. While international models incorporate a wider range of factors, this framework prioritizes core criteria by grouping them into three main categories–attack target, attacker intent, and impact–comprising a total of seven sub-criteria, all organized into three levels for simplicity.

Attack Target

1. *Scope:* This assesses scope by affected entity range.
    - High: Nationwide or multinational impact

- Medium: Regional or multi-organization impact
- Low: Localized or single-organization impact

2. *Importance of Affected Targets*: This assesses the criticality and sensitivity of the affected organizations and systems.

    a. High: Organizations crucial to national security, such as critical infrastructure or government agencies, based on the **critical infrastructure** sectors defined by the U.S. Department of Homeland Security (DHS) and the ROK's **Act on the Protection of Information and Communications Infrastructure**

    b. Medium: Large-scale private sector organizations, including major enterprises. The criteria for defining large enterprises in the ROK and the United States are as follows (though they may differ by industry in both countries):

        i. South Korea: Large-scale enterprises are defined by the Korean government as organizations with total assets exceeding **10 trillion KRW**.

        ii. United States: Large enterprises are generally defined as those with more than 500 employees or annual revenue exceeding **$47 million**, according to the Small Business Administration (SBA).

    c. Low: Smaller organizations or those with limited national significance

Attacker Intent

3. *Intent:* This assesses the goals and motivations of the attacker.

    a. High: Goal-oriented attackers—including state-sponsored groups or international hacking organizations—conducting long-term operations for political or economic gain

    b. Medium: Criminal actors pursuing monetary gains

    c. Low: Opportunistic or unintentional actors with no clear objectives

Impact

4. *Functional Impact:* This assesses the disruption or degradation of critical functions and services. Time thresholds, such as the recovery time objective (RTO) or maximum tolerable downtime (MTD), are used as benchmarks. The time thresholds may vary depending on the importance of the organization and system.

    a. High: Significant damage or prolonged interruptions, such as system downtimes exceeding the RTO or MTD

        i. South Korea: The RTO for critical government systems can be set within three hours. The Ministry of Interior and Safety in South Korea has set the MTD for government systems at **three hours**.

ii. United States: The RTO for national or mission-essential functions is generally set within 12 hours. According to **NIST Special Publication 800-34 Rev. 1** (Contingency Planning Guide for Federal Information Systems), the recovery time for functions that are national, primary, or mission-essential must be within 12 hours to ensure operational continuity.

a. Medium: Partial damage or interruptions, such as disruptions contained within the RTO or MTD limit

b. Low: Minor or no impact on functions and services

5. *Information Impact:* This assesses the compromise of data integrity, confidentiality, or availability:

   a. High: Breach of classified or sensitive information

   b. Medium: Breach of non-classified or general information

   c. Low: Minor or no impact on information

6. *Economic Impact:* This assesses the financial and economic damage caused by cyberattacks. The thresholds used are reference values informed by the U.S. **CISA report** *Cost of a Cyber Incident,* which systematically analyzes financial impacts (with a methodology detailed in Appendix A of this paper).

   a. High

      i. South Korea: Financial loss exceeding 20 billion KRW (approximately $15.7 million) or affected population of more than 10 million individuals

      ii. United States: Financial loss exceeding $157 million or an affected population of more than 72 million individuals

   b. Medium

      i. South Korea: Financial loss between 5 billion and 20 billion KRW (approximately $4 million to $15.7 million) or affected population of 1–10 million individuals

      ii. United States: Financial loss between $40 million and $157 million or an affected population of 8–72 million individuals

   c. Low

   i. South Korea: Financial loss below 5 billion KRW (approximately $4 million) or an affected population of fewer than 1 million individuals

   ii. United States: Financial loss below $40 million or an affected population of fewer than 8 million individuals

7. *Political and Social Impact:* This considers the implications for national security, reputation, public relations, and societal responses.

a. High: Severe impacts, including extensive media coverage, significant social media backlash, substantial drops in presidential approval ratings, or a decline in international trust indices
b. Medium: Moderate impacts with notable effects on public perception or international relations
c. Low: Minor or localized effects with limited public attention

WEIGHTING OF ASSESSMENT CRITERIA

According to the **Tallinn Manual 2.0**, a cyber operation's classification as a use of force or severe attack depends on scale, effects, and target. And the **UN GGE** and **OEWG**, two international initiatives on cyberspace norms, emphasize that attacks on critical infrastructure pose significant risks and are considered particularly severe.

A **survey** conducted by the OSCE identified impact and damage scale as the most critical criteria for assessing cyberattacks, followed by attack target. Similarly, a 2022 survey by the National Security Research Institute (NSR) in South Korea found that experts ranked attack sector, importance of the damaged system, and damage scale as the top factors influencing cyber severity assessments (detailed in Appendix B of this paper).

This prioritization aligns with a broader consensus that cyberattack assessments should focus on:

- *Impact,* including social, functional, and economic consequences
- *Target,* especially when critical infrastructure is affected

The CSCF reflects importance by letting the number of items in each category influence the results.

SCORING AND SEVERITY LEVELS

The CSCF determines severity by assigning scores to assessment criteria, with a maximum score of 35 points. Each criterion is scored on a three-tier scale based on severity: High (4–5 points), Medium (2–3 points), and Low (0–1 point). The framework uses the total score across all criteria to categorize cyber incidents into six severity levels—Normal, Low, Medium, High, Severe, and Critical—aligning with the CISS used in the United States and France. The United Kingdom and Australia also follow a similar six-tier classification model. Using comparable severity levels helps promote mutual understanding of assessment results among nations.

Table 2: Cyberattack Severity Level

Level	Score Range	Details
Normal (White)	0–10	No impact on national security or citizens
Low (Green)	10–15	Minimal impact on national security and citizens; does not affect major functions
Medium (Blue)	16–20	Impact on national security and citizens; includes partial disruption to major functions and resources
High (Yellow)	21–25	Significant impact on national security and citizens; affects multiple major functions and resources
Severe (Orange)	26–30	Severe impact on national security and citizens; includes widespread disruption to major functions and resources
Critical (Red)	31–35	Critical and nationwide impact on national security and citizens, causing extensive damage and disruptions

Source: Authors' analysis.

Cumulative Effects

Cyberattacks from major adversarial states are often parts of broader strategic **campaigns** rather than isolated incidents. These operations typically unfold within the "gray zone," staying below the threshold of armed conflict to avoid direct retaliation. Attackers employ long-term reconnaissance and persistent access, expanding their targets from individual systems to interconnected networks to achieve their strategic objectives.

Assessing cyber incidents in isolation risks underestimating their intent and cumulative effect. To address this, viewing a cyberattack as part of a broader campaign not only strengthens the assessor's ability to make **political attributions** but also provides a deeper understanding of the attacker's strategic objectives and operational patterns. In line with this, U.S. **PPD-41** recognizes that a series of related incidents with cumulative effects may qualify as a Significant Cyber Incident, and NATO's 2021 **Comprehensive Cyber Defence Policy** acknowledges that such activities could, under certain conditions, be classified as an armed attack.

Integrating cumulative effects into the severity assessment process also helps mitigate a key limitation of the CSCF: the potential perception that low-level or sub-threshold attacks will go unaddressed. By accounting for the cumulative effect of recurring or persistent malicious activity, the framework conveys that even seemingly minor incidents may be evaluated and incorporated into broader strategic assessments.

To evaluate cumulative effects, a composite approach should be applied across the assessment criteria. For each criterion, the scope should be aggregated across incidents; the importance of affected targets should be assessed based on the most critical target; intent should reflect the most hostile objective; and impacts should be accumulated to reflect the overall effect.

CASE STUDIES

This paper's case study applied the CSCF to 21 cyberattacks over the past decade, primarily targeting the ROK and United States, focusing on critical infrastructure and state-sponsored hacking activities. The results are presented in Table 3. While the initial assessment covered 21 cases, 2 incidents were re-evaluated by applying cumulative effect, bringing the total to 23 entries in the table. Entries 1 through 8 correspond to the ROK and 9 through 20 to the United States. Entry 21 involves a German target, included for cumulative analysis, while entries 22 and 23 reflect re-assessed cases.

The assessment presents **EuRepoC's** cyber intensity and impact scores as a comparative benchmark to provide context for the CSCF results. Where EuRepoC scores are unavailable, the corresponding table entries are marked as blank. **EuRepoC**, an open-access database supported by EU member state governments, measures cyber intensity based on direct effects and sociopolitical severity, while its impact indicator evaluates broader consequences.

EuRepoC's cyber intensity is categorized into three levels:

- 1–5: Low/Moderate (Green); 6–10: High (Yellow); 11–15: Very High (Red)

Impact is classified into five levels:

- 1–5: Minor (White); 6–10: Low (Green); 11–15: Medium (Yellow); 16–20: High (Orange); 21–25: Very High (Red)

In South Korea, cyberattacks ranged from Medium to High severity according to the CSCF, with cases like the KHNP Hacking (Entry 1) and the Pyeongchang Olympics Hacking (Entry 3) showing notable social and political impact. Although the KHNP Hacking caused limited direct disruption, it triggered public alarm by targeting nuclear facilities, and led to the development of **South Korea's cybersecurity strategy**. It was classified as High severity under the CSCF. However, according to EuRepoC, both cases were rated within the Low/Moderate intensity range.

North Korean hacking groups were responsible for the most severe attacks on South Korea, particularly the 2023 Andariel IT Company Breach (Entry 7) and the 2024 Defense Contractor Breach (Entry 8), both of which targeted military intelligence and defense technology. The Defense Contractor Breach received the highest scores under the CSCF due to its significant informational and social impact. In contrast, according to EuRepoC, Entry 7 was rated as Low/Moderate in intensity and Low in impact, while Entry 8 was assessed as High in intensity but still rated as Low in impact.

The United States, facing more frequent and severe cyber threats compared to South Korea, exhibited higher severity levels in CSCF-based assessments, with attacks classified as Medium to Severe. High-profile cases such as SolarWinds (Entry 13), Colonial Pipeline (Entry 17), Volt Typhoon (Entry 18), and Salt Typhoon (Entry 19) reached Severe under the CSCF due to their impact on

national security and critical infrastructure. Additionally, the Hive Ransomware Attack (Entry 14) was rated Severe according to the CSCF, since it caused significant financial and operational disruptions. In contrast, according to EuRepoC, the Hive Ransomware Attack was rated as Low/Moderate in intensity and Low in impact. For the remaining cases, while some were assessed as High in intensity, all were rated Medium or lower in impact.

Cases in which cumulative effect is applicable under the CSCF include the 2022 Defense Manufacturer Attack (Entry 6), and the 2024 Defense Contractor Breach (Entry 8), both by North Korean groups targeting South Korean defense contractors; and the 2024 Diehl Defense Breach (Entry 21) involving a German missile supplier, which also fits this pattern. Though individually limited, their cumulative effect within the "North Korea Defense Breach Campaign" justifies a reclassification to Severe.

The 2015–2016 DNC hacking incidents (Entry 10) in the United States are attributed to Cozy Bear and Fancy Bear as part of Russia's election interference efforts. Along with the post-election spear-phishing attacks (Entry 11) by the same threat actors and with the same political aim, these incidents can be collectively referred to as the "Russian U.S. Election Interference Campaign." When assessed as a broader campaign, the impact increases and may justify a Severe classification under the CSCF.

Overall, the CSCF and EuRepoC showed a similar relative ranking of cases—those rated highly under the CSCF generally also received higher ratings under EuRepoC. However, EuRepoC tended to assign lower absolute intensity and impact scores compared to the CSCF. This alignment in severity levels was slightly closer in U.S. cases, while ROK cases showed some differences, likely due to EuRepoC's EU-centric focus on large-scale impacts. Since political and economic impacts vary by national context, frameworks like the CSCF—which present reference points informed by such context—are better suited to national-level assessment. In particular, because the importance of the attack target is a critical factor in national security considerations, the CSCF has enabled more appropriate severity assessments in such contexts.

Table 3: Assesment Results for Cyberattack Cases

No.	Date	Name	Attributed Country	CSCF Score and Level	Notes (EuRepoC)	
					Intensity	Impact
1	Dec. 2014	KHNP Hacking	North Korea	21	3	-
2	June 2016	Cyber Command Hacking	North Korea	20	3	-
3	Feb. 2018	PyeongChang Olympic Hacking	Russia	21	4	-
4	Jan. 2019	Leak of Foreign Ministry Emails	China	18	4	-

5	May 2021	SNU Hospital Hacking	North Korea	18	4	6
6	Oct. 2022	North Korean Defense Manufacturer Attack	North Korea	23	4	7
7	Dec. 2023	Andariel's Attack on Korean IT Companies	North Korea	21	4	7
8	Jan. 2024	North Korean Defense Contractor Breach	North Korea	23	6	9
9	2014	Westinghouse Hacking	Russia	20	2	-
10	2015–2016	DNC Email Leak & Election Interference	Russia	23	4	-
11	Nov. 2016	Cozy Bear Post-Election Spear-Phishing	Russia	17	3	-
12	2017	Equifax Data Breach	China	21	4	-
13	2019–2020	SolarWinds Supply Chain Attack	Russia	28	4	15
14	2020	Hive Ransomware Attack	Non-state	26	4	7
15	2021–2023	RedHotel Attack	China	24	1	7
16	2021	Andariel Maui Ransomware Attack	North Korea	23	4	9
17	May 2021	Colonial Pipeline Attack	Russia	26	6	12
18	2021–2024	Volt Typhoon Attack	China	28	6	13
19	2023–2024	Salt Typhoon Attack	China	26	1	5
20	2023–2024	RedJuliett (Flax Typhoon) Cyberattack	China	25	3	7
21	Sept. 2024	Hacking of German Defense Company Diehl Defence	North Korea	15	-	-

22	2020–2024	North Korea Defense Breach Campaign	North Korea	29	-	-
23	2015–2016	Russian U.S. Election Interference Attack	Russia	27	-	-

Source: Authors' analysis.

Conclusion

The CSCF provides a systematic framework for assessing cyberattack severity, thereby enhancing situational awareness and supporting the ROK government's political decisionmaking. It promotes international understanding, facilitates national cost imposition efforts, and supports joint responses. This approach aligns with the strategic goals of the **ROK-U.S. Cybersecurity Cooperation Framework**, which emphasizes countering and deterring malicious cyber activities while holding responsible states accountable.

The CSCF was developed through comparative analysis to identify core criteria and organize them into three categories: impact, target scope, and intent. The framework also offers detailed sub-criteria with reference points based on publicly available data. When applied to cases, high-profile incidents such as the Volt Typhoon and Salt Typhoon were classified as Severe, effectively capturing their national security implications. Moreover, when multiple attacks by the same hacking group with similar objectives were assessed cumulatively as part of a single campaign—such as the North Korea Defense Breach Campaign—the attacks' severity levels increased.

The case study results underscore the need for a national-level classification framework that reflects country-specific contexts. The CSCF is particularly well-suited for evaluating the severity of cyberattacks in South Korea, especially from a national security perspective. Such a framework enhances understanding of domestic cyber incident severity and can serve as a foundation for improving mutual understanding in the context of future ROK-U.S. cybersecurity cooperation.

Cyberattacks have grown in scale and complexity over the past two decades. Unlike traditional military deterrence, deterrence in the cyber domain lacks clear red lines and established response thresholds, and allows a degree of anonymity. As a result, few cyberattacks have led to meaningful consequences, and major threat actors continue to exploit cyber capabilities. Therefore, active national-level responses that impose costs and ensure accountability are essential. Ultimately, response depends on national resolve, operational capacity, and coordinated action among like-minded nations. Nevertheless, continued research is needed not only to support political decisionmaking but also to foster international consensus. In addition, ensuring consistent assessments requires a dedicated agency for ongoing evaluations—as seen in the United States and United Kingdom—highlighting the need for a similar system in the ROK. Such an agency should share assessment results with stakeholders and foster a unified understanding of cyberattack severity.

Appendix A: Establishing Economic Impact Benchmarks

While the economic impact of cyber incidents is being studied in various fields such as cyber insurance, there is currently no established global standard. To address this gap, this study references the U.S. CISA's *Cost of a Cyber Incident* **report**, which systematically estimates national-level financial impacts using both commercial and research data. CISA analyzes 12 large incidents, categorizing costs as direct or indirect, and emphasizes relative costs—such as impacts compared to an organization's revenue or a nation's GDP—rather than absolute figures. However, the report acknowledges limitations, including the significant influence of outliers on total cost calculations.

Table A-1: Costs, Cost-to-Revenue Ratios, and People Affected (Large Incident Sample)

Company Affected	Year of Incident	Total Cost ($ million)	Cost-to-Revenue Ratio	Number of People Affected (millions)
Anthem	2015	375.5	0.48%	78.8
Yahoo	2014	350	7.58%	500
Merck	2017	310	0.78%	Unknown
Target	2013	292	0.41%	70
Home Depot	2014	252	0.30%	56
Sony PlayStation	2011	171	0.20%	101.6
Equifax	2017	164	4.88%	145.5
Sony Pictures	2014	43	0.06%	0.047
Experian	2015	20	0.42%	15
Yahoo	2014	16	0.34%	1000
Ashley Madison	2015	12.8	11.74%	37
LinkedIn	2012	4	0.41%	6.5

Source: Cybersecurity & Infrastructure Security Agency, *Cost of a Cyber Incident: Systematic Review and Cross-Validation* (Washington, DC: CISA, October 2020), https://www.cisa.gov/resources-tools/resources/cost-cyber-incident-systematic-review-and-cross-validation.

To analyze the economic costs and number of affected people in typical large-scale incidents while minimizing the influence of outliers, a 25 percent trimmed mean was applied. Based on this method, the trimmed mean for economic costs was calculated at $157 million, with the average number of victims at 71.99 million. Considering the approximate tenfold GDP difference and sevenfold population difference between the ROK and the United States, the threshold for large-scale incidents in the ROK can be estimated at approximately 20 billion KRW in economic

cost and 10 million people affected. Therefore, the suggested thresholds for assessing the economic impact of cyberattacks in South Korea—based on financial loss and the number of individuals affected—would be as follows:

- High: Over 20 billion KRW (approx. $15.7 million) and more than 10 million individuals affected
- Medium: 5–20 billion KRW ($4–15.7 million), affecting 1–10 million individuals
- Low: Less than 5 billion KRW (around $ 4 million) and fewer than 1 million individuals

Appendix B: Expert Survey on Assessment Criteria Importance

An expert assessment using the Analytic Hierarchy Process (AHP) and the Likert method was conducted from October 24–30, 2022, to evaluate the importance of cyberattack assessment items. Experts rated each item on a 1–9 scale; nine detailed evaluation items were considered. Eleven experts in cyber law, policy, and technology (with an average experience of 18 years) participated.

Among detailed items, "importance of the damaged system" and "attack sector" were rated highest, followed by "scale of damage," "informational impact," "social impact," and "attacker intent." "Functional impact" and "recoverability" ranked slightly lower, while "attack complexity/sophistication" received the lowest importance score. Although item rankings varied slightly, overall results consistently emphasized the significance of attack targets and damage scale over technical capability.

Table B-1: Assessment Criteria Importance

Number	Importance	Ranking
Importance of the damaged system	7.989	1
Attack sector	7.951	2
Scale of damage	7.535	3
Informational impact	7.526	4
Social impact	7.521	5
Attacker intent	6.797	6
Functional impact	6.781	7
Recoverability	6.749	8
Attack complexity/sophistication	5.772	9

Source: Authors' analysis.

Sunha Bae *is a senior researcher in the Cybersecurity Policy Department of the National Security Research Institute.*

Part II

Building Korea's Active Cyber Defense Strategy

Mutual Defense in Cyberspace

Joint Action on Attribution

By Julia Brock and James Andrew Lewis

On the seventieth anniversary of the U.S.-Republic of Korea (ROK) alliance, the leaders of the two countries **declared** that the Mutual Defense Treaty applies to cyberspace. The United States and South Korea agreed to expand the scope of the agreement and promote cooperation in cybersecurity technologies, policies, and strategies, including cyber threat information sharing.

The United States and the ROK have significantly strengthened their cybersecurity cooperation in recent years. South Korea's membership in NATO's Cooperative Cyber Defense Centre of Excellence laid the groundwork for enhanced information sharing, joint exercises, and the development of shared standards. The U.S.-ROK Cybersecurity Working Group, **established** in 2022, has focused on strengthening mechanisms for collaboration and joint defense, as well as defensive strategies. Both countries have also emphasized the importance of international partnerships in addressing cyber threats posed by authoritarian states. The U.S.-ROK Strategic Cybersecurity Cooperation Framework, signed in April 2023, further solidified this commitment by increasing information sharing and joint response efforts.

Washington and Seoul are actively engaged in countering cyber threats, particularly those from the Democratic People's Republic of Korea (DPRK). The Counter Ransomware Initiative (CRI), the U.S.-ROK-Japan Trilateral Summit, and the ROK's 2024 National Cybersecurity Strategy each

highlight the shared goal of disrupting the DPRK's malicious cyber activities and protecting critical infrastructure. Recent memoranda of understanding (MOUs) between the United States and the ROK have further strengthened cooperation in areas such as computer emergency response team **communications** and supply chain **resilience**.

Current Cooperation Structures

The overarching structure for the U.S.-ROK alliance is the **2023 Strategic Cooperation and Coordination Framework** (SCCF). The SCCF provides a diplomatic mechanism to facilitate discussions between the two countries on a range of issues, particularly those concerning security, military cooperation, and regional challenges. The SCCF aims to enhance bilateral coordination, improve strategic alignment, and address concerns related to the Korean Peninsula, North Korea, and other regional security matters.

The framework reflects the growing importance of cybersecurity and is an important component of the broader U.S.-ROK alliance. By working together, the two countries hope to deter cyberattacks, protect critical infrastructure, and promote a secure and stable cyberspace. The core areas of collaboration include the following:

1. **Intelligence Sharing:** This encompasses detailed information about North Korean cyber operations, emerging threats, attack patterns, and potential vulnerabilities. Washington and Seoul maintain secure channels for information exchange and conduct joint threat analysis. They share technical indicators, malware signatures, and attribution data in order to build a comprehensive understanding of the cyber threat landscape.
2. **Critical Infrastructure Protection:** The agreement establishes protocols for protecting energy grids, financial networks, transportation systems, and healthcare infrastructure. Both nations conduct vulnerability assessments, develop resilience strategies, and coordinate incident response plans. They share best practices for infrastructure security and collaborate on supply chain security.
3. **Cybercrime Cooperation:** Both countries share investigative resources, forensic capabilities, and information about criminal activities. They coordinate efforts to combat ransomware attacks, financial fraud, and data breaches. Joint investigation teams tackle cross-border cyber incidents, and streamlined processes facilitate rapid information exchange about emerging threats and criminal techniques. For example, the work of the CRI provides a collective approach for attribution and accountability.
4. **Joint Attribution:** Joint attribution—tracking and identifying the perpetrator of a cyberattack—by the United States and the ROK (and perhaps also including Japan, Australia, and other Five Eyes partners) may have an inhibiting effect on an opponent, particularly if linked to measures to create accountability, such as sanctions, indictment and other measures of retorsion.
5. **Capacity Building:** Collaborative programs improve both nations' cybersecurity capabilities. These include joint cyber defense exercises, personnel exchanges, and shared

training programs. Technical expertise is exchanged in areas such as malware analysis, incident response, and network defense. The countries also collaborate on research and development projects for new cybersecurity technologies.

6. **International Cooperation:** The United States and South Korea work together in international forums such as the United Nations to promote responsible state behavior in cyberspace beyond bilateral relationships, and should continue to do so. They also advocate for common security standards, coordinate positions on cyber governance issues, and support regional capacity-building efforts. A multilateral approach involving other like-minded nations, including Australia, Japan, and the United Kingdom, would reinforce these efforts. The ROK relationship with the NATO Cooperative Cyber Defense Centre of Excellence is one such avenue for cooperation.

The SCCF included implementation mechanisms through a Joint Cyber Coordination Committee that oversees progress and adapts strategies as needed. Senior-level dialogues ensure continuous alignment of objectives and approaches. Annual reviews assess the effectiveness of cooperation and identify areas for enhancement. Both countries can agree to conduct joint cyber exercises under the umbrella of the SCCF.

The SCCF also included provisions for addressing emerging technologies like artificial intelligence (AI), quantum computing, 5G/6G networks, and the **Internet of Things**. The agreement allows for changes to accommodate new threats and technological developments in the cyber domain, and emphasizes sustainable, long-term cooperation while maintaining the ability to respond rapidly to immediate threats. This comprehensive approach strengthens both nations' cybersecurity postures while contributing to broader regional stability in the Indo-Pacific. It also represents a significant step forward in establishing trusted partnerships for addressing evolving cyber challenges in an increasingly interconnected world.

Following the adoption of the SCCF, both nations are working to enhance their collaborative cyber defense capabilities. South Korea's 2024 **National Security Strategy** emphasizes developing its cyber workforce and strengthening international partnerships. This cooperation is especially crucial given North Korea's extensive cybercrime operations, which directly support the country's missile and nuclear programs. While bureaucratic differences have historically complicated U.S.-ROK cyber collaboration, both nations are committed to improving intelligence sharing and technical cooperation.

South Korea's Cyber Threat Landscape

South Korea is located at one of the most important intersections for geopolitical rivalry, and therefore is a target for cyberattacks. The country's primary adversaries in cyberspace include North Korea, China, and Russia.

NORTH KOREA

North Korea is South Korea's most dangerous adversary. Pyongyang has significantly improved its capabilities and expanded its cyber operations over the last decade, using cybercrime as a

major source of funding. UN reports indicate that North Korean hackers **conducted** almost 60 cyberattacks on cryptocurrency companies between 2017 and 2023, stealing around $3 billion in total, **including** $1.7 billion in 2022 alone. These cyber activities now account for about **half** of North Korea's foreign currency income and may **fund** as much as 40 percent of its weapons of mass destruction programs, making cybersecurity a strategic priority. DPRK hacker groups have targeted South Korean digital infrastructure, including satellite facilities, courts, and defense contractors.

North Korea primarily employs cyberattacks for specific, calculated purposes such as financial gain, intelligence gathering, or disrupting South Korean infrastructure. While the regime has demonstrated the ability to launch sophisticated attacks, these are typically targeted and aimed at causing disruption rather than widespread destruction. North Korea is likely to continue its current strategy of using cyberattacks as a tool of coercion and intimidation. By launching smaller-scale attacks, the regime can signal its capabilities and influence regional dynamics without risking a full-scale conflict. However, if South Korea, the United States, and other allies further improve their cybersecurity defenses, North Korea may find it increasingly difficult to obtain funds and achieve its objectives through cyber operations.

CHINA

China's cyber operations against the ROK are multifaceted and ongoing. While specific details about these operations are often classified, public reports and cybersecurity research **suggest** several key areas of focus, including espionage and intelligence gathering. Chinese state-sponsored hackers have targeted South Korean government agencies, military installations, and defense contractors to steal sensitive information related to national security, military capabilities, and diplomatic strategies. Chinese hackers have also targeted South Korean **companies**, particularly those in high-tech sectors like semiconductors and electronics, to steal intellectual property and trade secrets.

China's cyberattacks against the ROK threaten South Korea's national security, economic interests, and democratic values. State-backed hackers have used social media platforms to spread disinformation and propaganda aimed at influencing public opinion in South Korea. This includes spreading false narratives, promoting pro-China sentiments, and undermining trust in democratic institutions. China has developed sophisticated cyberwarfare capabilities, including the ability to launch large-scale cyberattacks that could disrupt critical infrastructure and government systems. The South Korean government and private sector must continue to invest in cybersecurity defenses to protect against these threats.

RUSSIA

Russia has also used cyber actions against the ROK, primarily for intelligence purposes. Russia is believed to have engaged in cyber espionage against South Korean government agencies, defense contractors, and other sensitive sectors. The goal is typically to gather intelligence related to national security, defense technologies, **arms sales**, and geopolitical **dynamics** involving North Korea (especially given North Korea's decision to send troops to support Russia's invasion of Ukraine). Russian cyber actors often use techniques such as phishing emails, malware, and other types of cyber intrusion to infiltrate South Korean networks.

Threats to the United States

The United States has also been targeted by North Korea, China, and Russia. Given these threats, both Washington and Seoul share a strong interest in countering these cyber operations to protect their infrastructure and impede Pyongyang's nuclear weapons program. Despite the real threat from North Korea, however, China poses the greatest cyber threat to the United States. China leads the world in espionage-related hacking against the United States. China is also active in other intelligence areas, such as the use of clandestine agents and satellites, but communications espionage is the centerpiece of Beijing's intelligence program. China has had major successes against the United States, most recently with an **operation** commonly known as Salt Typhoon. This is only the latest Chinese cyberattack, and it has affected more than two dozen countries.

Salt Typhoon should not be seen as an isolated incident, but as part of a larger Chinese campaign to systematically exploit global telecommunications networks. An earlier campaign commonly known as Volt Typhoon saw China pre-position malicious code on U.S. critical infrastructure networks. Salt Typhoon may have also been used in pre-positioning malicious code on telecommunications **networks**. Pre-positioning goes **beyond** espionage and is often a precursor to attack.

Additionally, China has constructed a broad global signals intelligence surveillance system, and it appears that the country has a comprehensive **strategy** for cyber espionage and communications intelligence. Its initial focus was on commercial and technological espionage, as well as conventional politico-military spying. In the last decade, Beijing has expanded its efforts in both scale and scope to include preparing for disruptive actions against critical infrastructure.

Unfortunately, the international cybersecurity situation is unlikely to improve anytime soon. Defensive measures alone will be inadequate to protect the national interests of both the United States and the ROK. This means that Seoul, working in collaboration with Washington, will need to develop stronger capabilities to support what it **calls** "offensive cyber defense." Attribution is the first step.

Attribution

As stated previously, cyber attribution is the process of tracking and identifying the perpetrator of a cyberattack. Raising specific instances with the state responsible and asking for explanation and cessation is a possible first step, albeit not with North Korea, which has refused to engage with the ROK. However, attribution for diplomatic purposes will be more effective when accompanied by persuasive evidence (and adequacy, discussed later) and when carried out by many nations. Attribution can be done in both private and public engagement.

Cyber attribution is seen as essential by many countries, and the difficulties of attributing cyberattacks are often exaggerated. The demand by political leaders for high accuracy in attribution can constitute a significant burden and can make it difficult to respond to malicious cyber actors. While attribution in cyberspace is challenging due to the ability of opponents to exploit the anonymity it affords, a combination of techniques can allow for accurate attribution. The SCCF does not explicitly outline specific procedures for cyber attribution, but it does provide a framework

for cooperation that can support attribution efforts. By sharing intelligence and information, collaborating on investigations, and taking advantage of both countries' technical capabilities, the United States and South Korea can work together to improve their ability to attribute cyberattacks.

Creating an Attribution Framework

Attribution is not primarily a matter of technical capabilities (although a lack of capacity can be an impediment). For international relations, political attribution (a decision by one government to assign responsibility for an act) is more important. Decisions on attribution are primarily political and require a sound foundation in intelligence and analysis; many countries prefer to avoid public attribution, and the current level of information sharing among states is **inadequate** to support a collective approach. Creating a **framework** of technical and factual attribution combined with the political decision to act would be one way to draw attention to the political requirements for action.

States can use the 2015 UN Group of Governmental Experts (GGE) **norms** as a framework for action. Reducing the number, scope, and risk of malicious cyber actions will require mechanisms for cooperation and common understandings of attribution, proportionality, and managing any risk from responsive action. One norm of the UN Open-Ended Working Group (OEWG) **states** that "in case of [information and communication technology (ICT)] incidents, States should consider all relevant information, including the larger context of the event, the challenges of attribution in the ICT environment and the nature and extent of the consequences." The document also notes that attribution "is a complex undertaking, and a broad range of factors should be considered before establishing the source of an ICT incident." This approach can guide thinking on situational awareness and increased accountability.

DEFINING CREDIBLE ATTRIBUTION

Using legal precedent complicates and confuses the discussion of attribution. Political attribution is in the domain of sovereign states, not the courts. The evidentiary standards are different and incompatible. Most importantly, political attribution by states does not involve identifying the culpable individual beyond a shadow of a doubt. It involves identifying the state responsible for the action, or from whose territory the action originated.

Identifying the responsible state is the central element for decisionmaking in political attribution. Political attribution builds on the responsibility **agreed** upon by all UN member states to observe their commitment to ensure that malicious actions do not emanate from their territory, to cooperate with a victim state when asked for help, and to ensure they take action against the malicious actor. If that malicious actor is unable or unwilling to take action, the victim state is permitted to take action itself (consistent with international law), either collectively or individually.

Credible attribution and proportionality in any response based on that attribution are essential ingredients for a politically acceptable response to malicious cyber action. The factors that states should consider in attributing an attack and in using attribution as a tool to increase accountability and stability in cyberspace include the following:

- precedent (e.g., previous attacks);
- technical indicators (and having best practices such as logging in place before an incident);
- the target (criminals are unlikely to go after military targets);
- probable intent;
- effect (i.e., what data was exfiltrated, what services were disrupted);
- external sources of information (e.g., allies or the private sector); and
- supporting intelligence from human or technical sources.

States can use multilateral, regional, bilateral, and multi-stakeholder platforms to share best practices and information on the attribution of different types of ICT threats and incidents. Coordinated attribution of malicious activity will require better information sharing between partners and, perhaps, new mechanisms for sharing and harmonization. These include evidentiary standards and information-sharing mechanisms for coordination of any collective attribution.

Attribution in the context of accountability is not primarily technical. It is primarily political and requires a sound basis of intelligence for political decisions. However, as stated above, the current level of information sharing among states is inadequate. Creating a framework of technical and factual attribution, combined with the political decision to act, would be beneficial. Other factors help determine the degree of rigor required for attribution, such as whether the attribution is tied to a responding action, or whether the conclusion of an effort at attribution will remain internal or be made public.

It is not necessary to identify the individual responsible for a malicious cyber action. It is only necessary to identify from whose territory the attack emanated. It is, of course, valuable and reassuring to identify the individuals responsible, but this can make the task more difficult and is not necessary for political attribution. The fundamental point is state responsibility for cyber actions taken from their territory. This level of attribution does not satisfy the requirements of a court, but courts have little or no jurisdiction in a conflict between countries.

Two objections to this approach are (1) the risk of "false flag" operations and (2) the need for sufficient evidence to persuade a public audience. The first is overstated. Few false flag operations can withstand scrutiny. And while it may be more satisfying for a public audience to identify an individual culprit, it is no more necessary than identifying the pilot who flies over a border.

The desire for a high degree of certainty in attribution before taking any action reflects exaggerated concerns over the potential risk of escalation and a desire to avoid unintended consequences. It can also cause unnecessary delays. Escalation risk from attribution is also generally overstated—there has been no incident of escalation in the 30-year history of cyberattacks. The risk is manageable using the tools of **diplomacy**.

CHALLENGES TO JOINT ATTRIBUTION

There are several significant challenges in conducting joint cyber attributions. Technical asymmetries create coordination difficulties. The United States and South Korea have different

technical capabilities, tools, and methodologies for attribution. While both countries have sophisticated cyber capabilities, their systems and approaches may not always align seamlessly, potentially leading to gaps or inconsistencies in attribution analysis.

Classification and information sharing structures can affect collaboration. Each country has their own national security classification systems and restrictions on sharing sensitive intelligence. This can limit the depth and speed of information exchange needed for comprehensive attribution. Sometimes critical technical indicators or intelligence sources cannot be fully shared due to classification concerns. Different legal structures affect how evidence can be collected and used. The United States and South Korea operate under distinct legal systems with different standards for digital evidence, privacy protections, and admissibility requirements, which complicates efforts to build legally sound attribution cases that would hold up in both jurisdictions.

Political considerations sometimes create divergent priorities. While both countries share concerns about North Korean cyber activities, they may have different diplomatic sensitivities about attributing attacks to other state actors, particularly China. South Korea's geographic proximity and economic ties to China may influence its willingness to publicly attribute attacks.

Operational security risks increase with joint attribution. Coordination between two countries inherently increases the number of people and systems involved, creating more potential points of compromise. This can make it harder to maintain operational security during sensitive attribution investigations. The United States and South Korea might have different timelines and thresholds for when they feel confident enough to make attribution claims. One partner might prefer faster public attribution while the other wants more conclusive evidence before making statements.

As stated previously, attribution is the first step in an offensive cyber response by defenders—the term mentioned in the ROK's 2024 cyber strategy —and involves several considerations that must be carefully weighed. These considerations include navigating the political and legal complications surrounding attribution claims, relying on appropriate evidentiary standards to make credible attributions (standards that differ from those used in courts and the legal system), and determining both the accountability of actors and whether a response is warranted (along with defining what constitutes proportionality). Additionally, organizations must consider their target audience when making attribution claims, and whether that audience is technical, public, political, or diplomatic in nature. There are also important questions about accountability and an appropriate role of and reliance on the private sector in attribution efforts.

Joint attribution requires significant technical resources and personnel from both countries. Differences in available resources or competing national priorities can affect the depth and sustainability of collaborative attribution efforts, and any effort to create accountability. When undertaking joint attribution efforts, there are three key considerations: technical, legal, and political. The mechanics of information sharing pathways need to be formally defined. The underlying intelligence infrastructure also requires joint assessment—including data collection systems, threat intelligence platforms, forensics tools, analytical capabilities, and communication channels—and the levels at which collaboration should occur must all be identified. Organizations

must also determine the appropriate degree of transparency and capabilities sharing between the different parties involved in making attribution claims.

Joint attribution would be the gold standard and the foundation for states to implement and create accountability. The United States faces several significant challenges in conducting joint cyber attribution with South Korea, despite the two countries' long alliance and shared security interests. One of the main challenges is the disparate development of technical capacity across the world to conduct the necessary analysis of incidents to determine the culprit. Besides support for technical development, a key consideration is the need for trust and transparency for joint attribution. Given the sensitive nature of the information, governments have been reluctant to share details of their technical attribution capabilities.

Attribution for creating accountability is not the same as attribution required by a court. In international affairs, attribution must provide sufficient information to persuade decision makers and both domestic and global audiences on the source of a malicious cyberattack. An overly legalistic approach cedes advantage to opponents. This is not a criminal proceeding, and it is essential to recognize that political attribution involves assessing the culpability of a state, not an individual. Attribution is difficult, but after decades of hostile action, waiting for more data to justify a response would be irresponsible. The difficulty of attribution should not be an excuse for inaction.

Attribution creates the conditions needed to validate an action in response to a malicious cyber act. This does not require identifying the individuals responsible, but rather the state responsible for the attack or for failing to observe its obligations under international law. Credible attribution and proportionality in any response based on that attribution are essential ingredients for a politically acceptable response to malicious cyber action.

KEY CRITERIA FOR JOINT ATTRIBUTION

The mechanics of information sharing pathways for attribution need to be formally defined for successful joint attribution. The underlying intelligence infrastructure also requires assessment, including data collection systems, threat intelligence platforms, forensics tools, analytical capabilities, and communication channels. Organizations must also determine the appropriate degree of transparency and capabilities sharing between the different parties involved in making attribution claims. Criteria for attribution include the following:

1. identifying the likely violators of South Korean or U.S. sovereignty, judging by their public actions and statements;
2. asking whose strategic interests are served by a cyberattack and violating sovereignty (this can be determined from publicly available information, internal assessments, and consultations with allies and partners);
3. previous incidents pointing to a particular state as the responsible actor;
4. similar or simultaneous violations in other states with attributed sources;
5. allies or friendly nations supplying supporting information;
6. evidence that an incident is part of a larger campaign;

7. technical indicators or other intelligence that points to a perpetrator (e.g., using information from a national source, a commercial firm, or an ally or partner); and
8. accurate past assessments of attribution by the relevant national services.

There is still a tendency in cybersecurity to overvalue technical aspects of attribution. This is not the kind of attribution required by a court of law. There are no judges in cyberspace, impartial or otherwise, and the evidentiary standards required in court are profoundly inappropriate for relations among states. Attribution of the source of a malicious cyber act will remain a national decision and any agreement on collective action must recognize this. A sovereign state has the right to decide who has attacked it. States will not give up that right. Attribution remains the prerogative of states, and a collective response will depend on agreement among states.

Recommendations for an Offensive Cyber Defense

Offensive cyber defense has four elements: adequate attribution capabilities, a menu of proportional responses, a framework for collective action, and the political will to act. It is also useful to find ways to engage in diplomacy with opponents directly (if not always publicly). South Korea has shifted its cybersecurity strategy to offensive cyber defense. This has involved moving from a primarily defensive posture to a more proactive approach in which the country actively seeks to identify and counter cyber threats before they materialize. This method utilizes offensive capabilities to disrupt potential attackers, particularly in response to threats from North Korea.

Offensive cyber defense should have two objectives. The first is to reduce malicious activity by opponents. The second is to create incentives for opponents to come to the negotiating table. At the moment, China, Russia, Iran, and North Korea have no incentives to either stop their attacks or negotiate. It has been a decade since the last serious cyber talks between potential belligerents, and offensive cyber defense works best if it is a central part of a larger diplomatic strategy engaging allies, third countries, and, ideally, opponents. No cyber defense will be adequate against determined, well-resourced, and inventive adversaries.

Offensive cyber defense is outlined in South Korea's 2024 National Cybersecurity Strategy, emphasizing the importance of attribution and retaliation against malicious actors. The ROK's Cyber Command plays a central role in this strategy. Established in 2010, the command oversees both defensive and offensive cyber operations. It is tasked with protecting the nation's digital infrastructure, defending against foreign cyberattacks, and, when necessary, executing offensive cyber operations in response to threats.

Offensive cyber defense includes proactive measures designed not only to protect and defend critical systems from cyberattacks—measures that rely on law enforcement or financial actions—but also to deter or neutralize potential threats before they can cause harm. This concept is part of a broader strategy to strengthen the ROK's cybersecurity posture and resilience. Instead of simply waiting for a cyberattack to occur, the ROK can identify threats early and act to prevent or mitigate them. Monitoring cyber activity, identifying vulnerabilities, and responding quickly to malicious actions will all be necessary to preventing cyberattacks. Offensive cyber defense includes the ability

to conduct counterattacks on adversary networks, either to disrupt or disable their operations or to prevent further escalation of cyber threats. Examples include hacking back, disrupting the attackers' infrastructure, or causing damage to their systems. Below are four strategies the ROK could consider for its offensive cyber defense:

1. **Coordinate offensive cyber defense operations with key allies, including the United States.** As part of the U.S.-ROK military alliance, the two nations share cyber defense capabilities and intelligence. Joint cyber exercises and information sharing help both nations strengthen their cyber resilience and readiness for offensive operations. The ROK has invested heavily in building sophisticated cyber tools, which may include malware, denial-of-service capabilities, and other advanced technologies designed to disrupt enemy networks. These tools are used to target adversary infrastructure and provide a strategic advantage in the event of conflict.
2. **Create clear legal frameworks to govern offensive cyber defense actions.** In practice, this involves ensuring that offensive operations do not violate international law or provoke unintended escalations. The ROK has been working to define the rules of engagement in cyberspace, making sure that its actions comply with national laws, international norms, and human rights principles, and this would be useful area for further consultation with Washington. One goal of offensive cyber defense should be to change opponent behavior by creating accountability. Absent accountability and consequences for malicious cyber actions, these will continue to increase and will become increasingly destabilizing. This does not have to involve military action or offensive cyberoperations, although these should not be ruled out. There is no accountability in cyberspace, but accountability has always been difficult for the international community.
3. **Develop both capabilities and a willingness to use them.** Even the United States, which has advanced cyber capabilities, has been reluctant to use the full range of these tools, despite the country's **policies** of active defense and "defend forward." While a few other Western nations are exploring the use of offensive cyber operations to create consequences, these operations have been few and have lacked noticeable effect. One reason for the lack of effect may be that responses tend to be episodic one-offs, rather than a set of sustained campaigns. This points to one central issue for creating accountability: Is this a response to an individual action, as would be the case in a law enforcement approach, or is it a response to a sustained opponent campaign? Policymakers have tried the former without success.
4. **Create a menu of response options and proactive measures.** Instead of solely trying to defend or prevent cyberattacks, South Korea aims to actively disrupt cyber operations by identifying vulnerabilities in potential attackers' infrastructure and taking preemptive actions. International cooperation is key, as South Korea also plans to strengthen collaborations with other countries to enhance its offensive cyber capabilities and intelligence sharing, with a focus on attribution.

Next Steps

The ROK revised its National Cybersecurity Strategy in 2024 as part of its active response to evolving economic, technological, and cybersecurity challenges. This revised strategy, which is aligned with UN Group of Governmental Experts requirements, focuses on transparency in cyber policies while ensuring safety and responsible use of cyberspace. The strategy encompasses three main features: (1) clear identification of cyber threats with a more proactive response approach, (2) reinforcement of the ROK's role as a "**Global Pivotal State**" in international cybersecurity cooperation, and (3) establishment of a comprehensive domestic cybersecurity governance framework.

To implement this strategy effectively, the ROK has outlined four crucial next steps:

1. developing detailed implementation plans;
2. communicating the nation's position clearly to the international community;
3. gathering objective data on partner countries' cybersecurity capabilities; and
4. establishing theoretical and practical foundations for more aggressive strategic approaches.

The United States can play a major role in helping South Korea implement its new cybersecurity strategy by offering technical support, sharing best practices, and fostering bilateral cooperation on cybersecurity issues. Washington can assist Seoul in enhancing its cybersecurity infrastructure by providing expertise in advanced technologies such as AI, machine learning, and threat detection systems. The United States, including U.S. tech companies, can collaborate with South Korean counterparts to implement cutting-edge tools for identifying and mitigating cyber threats from state actors including North Korea, China, and Russia.

The United States can also facilitate information sharing between South Korean and U.S. intelligence agencies. By exchanging threat intelligence data on emerging cyber threats and tactics used by adversaries, both countries can develop a more robust defense posture. The United States also has experience in addressing large-scale cyberattacks, such as those targeting critical infrastructure, which South Korea can learn from to bolster its own national resilience. This can include joint cyber exercises, technical seminars, and internships to help South Korea cultivate skilled professionals capable of handling complex cyber challenges.

A potential research agenda for U.S.-ROK cybersecurity cooperation can focus on several key areas, starting with threat intelligence sharing and analysis. Research questions in this area include how to improve the timeliness and accuracy of threat intelligence—especially regarding emerging risks like AI-powered attacks—and identifying effective mechanisms for joint threat analysis and attribution of cyberattacks from both state and non-state actors.

The next focus is on critical infrastructure protection, particularly in sectors like energy, finance, and transportation. Research could explore how to develop a joint framework for securing these sectors through information sharing, joint exercises, and technology cooperation. Additionally, strategies for enhancing the cybersecurity resilience of emerging technologies such as 5G within critical infrastructure should be examined.

The agenda could also emphasize cyber defense technology cooperation, seeking ways for the United States and ROK to collaborate on cutting-edge cybersecurity technologies including AI for threat detection, blockchain for secure data sharing, and quantum-resistant cryptography. Further, it could call for research on the legal and regulatory frameworks needed to facilitate cross-border technology cooperation while ensuring data privacy and national security. Other areas for possible collaboration include building human resources through joint education and training programs, as well as fostering international cooperation to establish norms for responsible state behavior in cyberspace.

Julia V. Brock *is a former program manager and research associate for the Strategic Technologies Program at the Center for Strategic and International Studies (CSIS) in Washington, D.C.* ***James A. Lewis*** *is a non-resident senior adviser in the Economic Security and Technology Department at CSIS.*

Forging Forward

South Korea's Proactive Cyber Defense and Strategic Cooperation with the United States

By Joohui Park and Donghee Kim

Introduction

When countries design their cyber defense strategies, they do so on the premise that they can defend against malicious cyber operations only after the impacts of such acts materialize. Instead of taking this reactive posture, however, the United States spearheaded a paradigm shift to a proactive posture in 2018, when the Department of Defense (DOD) developed the **"Defend Forward" strategy**. This posture has since been maintained in the United States, and was reaffirmed in the introduction of the **2023 Department of Defense Cyber Strategy**. Variations on this approach have been adopted by some countries, including the Republic of Korea (ROK). South Korea introduced a proactive approach to cyber defense in its 2024 National Cybersecurity Strategy, some aspects of which **mirror** the DOD's Defend Forward.

Proactive cyber defense is just one option available to countries responding to cyber threats. In general, response options are various and can be structured with a range of typologies. For instance, options can diverge into defensive and offensive responses; defensive responses, in turn, can be classified as either reactive or proactive. Public attribution and countermeasures under international law are examples of reactive defense; proactive defense can include active cyber

defense and other threat-hunting activities. In certain circumstances, alternatives to proactive cyber defense can be more effective. This paper is based on this conceptualization.

For proactive defense, countries need to glean insights into adversaries' acts in cyberspace before their impacts reach the intended targets. These insights are mostly embedded in physical infrastructures that adversaries exploit for malicious cyber operations—facilities primarily located in foreign territories. That is why working with allies and partners is crucial for proactive cyber defense. Fortunately, South Korea and the United States have made a great effort to cooperate in this space over the last few years. Such cooperation is primarily based on the alliance between the two countries. This paper aims to explore areas in which South Korea and the United States must deepen their collaboration to bolster their proactiveness in cyber defense.

South Korea's Cyber Defense Shift

MOVING TOWARD A PROACTIVE POSTURE

2019 National Cybersecurity Strategy

In 2019, the administration of former South Korean President Moon Jae-in unveiled South Korea's first **National Cybersecurity Strategy**. The defensive stance outlined in the 2019 strategy can be characterized as reactive rather than proactive. The strategy's second strategic task, "enhancement of cyberattack response capability," offers a glimpse into this reactive stance. According to the **2019 National Cybersecurity Basic Plan**, an implementation plan of the 2019 strategy, the administration, with a view to enforce cyberattack response capabilities, undertakes to (i) secure cyber deterrence by managing vulnerabilities and taking steps for attribution; (ii) fortify readiness against massive cyberattacks by revamping the framework for national response and developing cyberattack detection technologies; (iii) seek comprehensive and active means of response through international cooperation with like-minded partners and reinforcement of cyber warfare capacities; and (iv) improve cybercrime response by enhancing investigation and prosecution capabilities. In sum, the 2019 strategy appears to have been crafted without consideration of the proactive defense concept. The 2019 strategy is significant in that it represented the first-ever national cybersecurity strategy in South Korea and established a governance framework where the National Security Office of the Blue House took the lead in national cybersecurity. However, its strategic ideas concerning cyber defense seem to have remained at a rudimentary stage.

2024 National Cybersecurity Strategy

In February 2024, then-President Yoon Suk Yeol's administration made public the **second National Cybersecurity Strategy**, which shifted South Korea's approach to an offensive posture. In the 2024 strategy, the country set the "development of offensive cyber defense and response capabilities" as one of three primary objectives crucial to fulfilling its strategic vision of acting as a "global pivotal state." To that end, South Korea set forth five strategic tasks, one of which is "enhancing offensive cyber defense activities."

Although the 2024 strategy calls its defensive posture "offensive cyber defense," the strategy's posture is not so offensive that such naming is necessary, particularly considering the subtasks

that South Korea is indeed planning to carry out to fortify "offensive cyber defense." The strategy's subtasks are specified in the **National Cybersecurity Basic Plan**, which details the activities that the administration intends to implement in pursuit of offensive cyber defense. These include establishing public attribution; tracing threat actors' cyber infrastructure; issuing joint advisories; and collecting and analyzing threat intelligence over attack origins, among other things. Although some of these subtasks may have offensive elements, these activities do not comfortably fit under the term "offensive."

Furthermore, inadequate naming could give rise to mischaracterization and misinterpretation, and thus escalation. It is hard to draw a clear line between the defensive and the offensive in cyberspace. In the spectrum of cyber activities, purely defensive actions such as firewall protections sit at one end, while offensive operations to destruct or destroy targets sit at the other end. It is hard to tell where on this spectrum the administration's "offensive cyber defense" sits. Given this, use of the term "offensive" may risk signaling that South Korea would resort to the spectrum's offensive extreme, which is not the intention reflected in the strategy.

In order to enhance offensive cyber defense, the Yoon administration's **strategy** highlights three main activities, some of which can be regarded as reflecting a proactive approach. The first is attribution. In the strategy, South Korea stated that it will identify the perpetrators of cyberattacks that impair national security and national interests by mobilizing the legal and technical capabilities necessary for attribution. The country also emphasized that it will use scientific evidence to identify actors behind cyberattacks and hold them accountable for their malicious behavior. To this end, South Korea is set to establish **a procedure and standard for attribution** and employ it to build a foundation for international cooperation.

The second activity highlighted in the strategy is the strengthening of joint deterrence with partner countries. The administration seeks to maximize deterrence of threat actors through the issuance of joint cybersecurity advisories, which South Korea has actively issued with partners over the last few years. For instance, in July 2024, South Korea issued **a joint cybersecurity advisory on APT40**–a group sponsored by China–with the relevant ministries of Australia, Canada, Germany, Japan, New Zealand, the United Kingdom, and the United States. In the same month, South Korea, the United Kingdom, and the United States issued **a joint security advisory on the activities of the Andariel**, a hacking group under the Reconnaissance General Bureau of North Korea.

The third activity is preemptive and proactive response through active detection and attack origin analysis. This reflects an approach similar to the United States' Defend Forward. South Korea's strategy gives its intelligence agency and its military a preemptive and proactive mission to detect and analyze the sources of cyberattacks, catch signs of attacks in advance, and quickly share information with relevant ministries. To implement this, the **2024 National Cybersecurity Basic Plan** states that South Korea will develop technologies to identify threat actors and track their bases, infrastructure, and activities. Notably, there is a nuanced difference between attribution and the identification of the source of a cyberattack. The former is a procedure necessary to impose responsibility–i.e., legal or political costs–on persons or states conducting malicious acts in cyberspace. On the other hand, the purpose of source identification is to proactively track

down the source of the attack before the intended target is affected, reducing the impact (if any) on the target's network. With this third element, South Korea can be seen as taking a step toward proactive cyber defense.

ORGANIZATIONAL FRAMEWORK AND PROACTIVE CYBER DEFENSE

In South Korea, the National Security Office (NSO) is **responsible** for coordinating overall cybersecurity-related tasks as well as setting up and reviewing mid- to long-term policy directions. Under the coordination of the NSO, governmental agencies and departments carry out their respective cybersecurity-related responsibilities. In particular, the National Intelligence Service (NIS) and the Cyber Command under the Ministry of National Defense (MND) play central roles in proactive cyber defense.

As South Korea rearranged its cybersecurity strategy in 2024, the National Intelligence Service Act and its presidential decree, the Regulation on Cybersecurity Duty, were amended accordingly and entered into force. As a result, the NIS's four major cybersecurity duties became more refined and streamlined. First, the NIS performs intelligence activities for national cybersecurity. That is, the NIS collects, prepares, and distributes intelligence on international hacking organizations or state-sponsored hacking groups (**NIS Act**, Article 4(1)(1)(e); **Regulation on Cybersecurity Duty**, Article 3(1)(a)). Second, the NIS takes measures to identify, deter, and block threatening cyber activities conducted by North Korea; foreign states, nationals, or organizations; transnational actors; or South Koreans affiliated with any of these entities (NIS Act, Article 4(1)(3); Regulation on Cybersecurity Duty, Article 3(1)(b)). Third, the NIS takes preventive and responsive measures against cyberattacks and threats on public entities (NIS Act, Article 4(1)(4); Regulation on Cybersecurity Duty, Article 3(2)(a)). Fourth, the NIS is entitled to establish and operate a consolidated response governance in which public and private entities jointly work to manage and respond to crises (Regulation on Cybersecurity Duty, Article 6 bis (4)). In the event of a crisis, the NIS's National Cyber Security Center (NCSC) serves as the Cyber Threat Task Force (CTTF) tasked with responding; in times when there is no active crisis, the agency functions as the National Cyber Risk Management Unit (NCRMU).

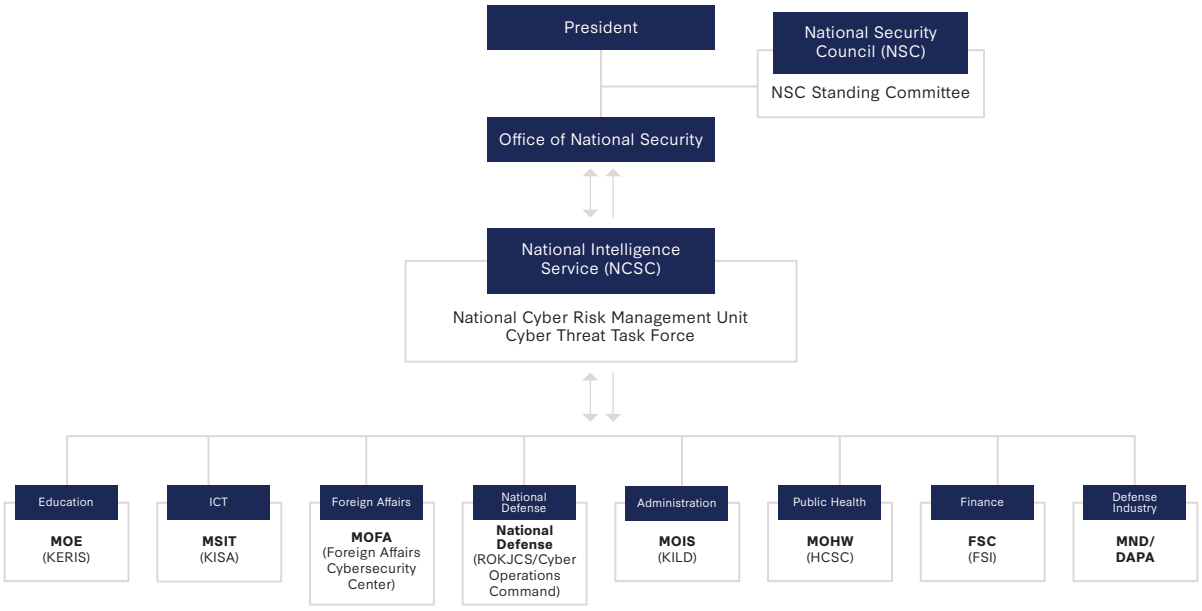
In the context of South Korea's proactive cyber defense, Article 6 bis of the Regulation on Cybersecurity Duty must be highlighted. This article was newly introduced in the revised Regulation on Cybersecurity Duty in 2024. According to paragraph three of Article 6 bis, the director of the NIS may take necessary steps to proactively identify, deter, and block activities against national security and interests. Such measures under this provision may include tracking and neutralizing foreign and North Korean bases. This provision provides a legal basis for the NIS's proactive cyber defense.

South Korea's Cyber Command also assumes cybersecurity tasks, mostly those related to national defense. The Cyber Command was established under the MND to take control of these duties in 2011 (**Presidential Decree on Cyber Command**, Article 1; the Cyber Command was originally established in 2010 under the Defense Intelligence Command of the Ministry of National Defense. In 2011, the Cyber Command became a unit directly subordinate to the Ministry of National Defense, per the Presidential Decree on Cyber Command enacted in that year). The Cyber Command's duties include planning and executing cyber operations and related cybersecurity needs; developing and

establishing frameworks necessary for cyber operations; and collecting, analyzing, and utilizing cyber threat intelligence (Presidential Decree on Cyber Command, Article 2). In addition, the Defense Counterintelligence Command under the MND supports cyber defense and information warfare (**Presidential Decree on the Defense Counterintelligence Command**, Article 4(5)).

South Korea's governmental structure for cybersecurity is illustrated in the graphic below.

Figure 1: South Korea's National Cybersecurity Implementation Framework



Note: MOE refers to the Ministry of Education; KERIS refers to the Korea Education and Research Information Service; MSIT refers to the Ministry of Science and ICT; KISA refers to the Korea Internet and Security Agency; MOFA refers to the Ministry of Foreign Affairs; ROK JCS refers to the ROK Joint Chiefs of Staff; MOIS refers to the Ministry of the Interior and Safety; KILD refers to the Korea Local Information Research and Development Institute; MOHW refers to the Ministry of Health and Welfare; HCSC refers to the Health and Welfare Cyber Security Center; FSC refers to the Financial Services Commission; FSI refers to the Financial Security Institute; MND refers to the Ministry of National Defense; and DAPA refers to the Defense Acquisition Program Administration.

Source: 2024 National Cybersecurity White Paper.

South Korea's Cooperation with the United States for Proactive Cyber Defense

THE U.S. APPROACH AND DEFEND FORWARD

The crux of proactiveness in the United States' cyber defense is reflected in the "Defend Forward" posture. In 2018, the year after President Donald Trump began his first term in office, the DOD's **cyber strategy** introduced Defend Forward, which represented a new approach: the idea of moving as close as possible to the origin of an adversary's activity as a way of defending against it. Defend Forward was first introduced in the 2018 vision document of the U.S. Cyber Command:

Achieve and Maintain Cyberspace Superiority: Command Vision for US Cyber Command. According to the document, the United States will defend forward to disrupt and stop malicious cyber operations—including activities below the level of armed conflict—at their origin. The United States emphasizes that although Defend Forward involves activities on the adversary's network and not the U.S. network, **this is still a defensive activity**, not an offensive one.

Defend Forward is inextricably linked to the concept of persistent engagement. The United States introduced the philosophy of persistent engagement in recognition of the fact that Cyber Command needs to engage with enemies on an ongoing basis in order to disrupt or impair their capabilities. The goal of persistent engagement is to identify and stop overseas cyber threats before they reach the U.S. network. Persistent engagement is designed to ensure that Cyber Command's Cyber National Mission Force takes and maintains the initiative to succeed in **the daily competition** with enemies.

The spirit of Defend Forward and persistent engagement is reflected in a **metaphor** from General Paul Nakasone, former commander of U.S. Cyber Command: "Our naval forces do not defend by staying in port, and our airpower does not remain at airfields. They patrol the seas and skies to ensure they are positioned to defend our country before our borders are crossed. The same logic applies in cyberspace."

The Biden administration continuously adopted Defend Forward under Pillar II ("Disrupt and Dismantle Threat Actors") of its **2023 National Cybersecurity Strategy**. The strategic tasks in Pillar II aim to mobilize all powers to prevent malicious cyber actors from threatening national security and public safety. The United States elaborated that it would expand private sector participation in neutralizing malicious cyber activities and promote cooperation with international partners. The **DOD** signaled that, in line with the 2023 National Cybersecurity Strategy, it will continuously defend forward by disrupting the malicious cyber activities and degrading their supporting ecosystems. (In the same year, General Nakasone **stated** that "There was a huge inflection point in 2018 with the Defend Forward [strategy]. I don't see, necessarily, a huge change in the strategy coming out.")

ROK AND U.S. CYBER DEFENSE APPROACHES: DIFFERENCES AND COMMONALITIES

There are both shared features and divergences between the proactive cyber defense approaches of the ROK and the United States. In order to identify possible areas for cooperation between the two countries, these commonalities and differences need to be analyzed.

Differences

First, the U.S. and South Korean approaches have different backgrounds. One of the underlying reasons for the United States' paradigm shift toward a proactive posture is the **failure** of the U.S. cyber deterrence strategy prior to 2018. In early cybersecurity strategy formulations, the United States primarily perceived acts that reached the use-of-force threshold as severe cyber threats. However, the malicious cyber operations to that point—such as the 2014 Sony Hack, the OPM breach, and the DNC hack—had been below the level of use of force. Minor but frequently occurring cyber incidents, meanwhile, had been neglected but were having cumulative and erosive impacts

on national security. Recognizing this threat landscape, the United States devised the strategies of **Defend Forward** and persistent engagement in order to contest daily competition in cyberspace. In contrast, South Korea's embrace of a proactive stance is not rooted in such reflection. Rather, South Korea took note of the inherent nature of cyber threats. In short, the hyper-connectedness of cyberspace makes complete prevention and defense against cyber threats too constrained.

Whatever the origin of South Korea's approach, it is evident that the major threats the country has experienced are not at or above the level of use of force. For instance, in May 2024, one terabyte of data from South Korea's Supreme Court was stolen by Lazarus, North Korea's hacking group. Yet physical and tangible damage—which would normally be calculated to assess whether an incident justifies use of force—was not caused. However, significantly sensitive personal information was **leaked**, including information on resident registration, marriage, and medical certificates. These types of damage could have a knock-on impact that could threaten national security. Accordingly, South Korea's cyber defense posture must be calibrated to adequately disrupt such low-level activities.

Second, leadership frameworks for defending proactively also differ between the two countries. In the United States, a single leader serves as the director of the National Security Agency (NSA) and the commander of the U.S. Cyber Command. This dual-hatted leader coordinates Defend Forward operations by integrating intelligence. However, the roles of the South Korean NIS and Cyber Command are not integrated for the purpose of proactive cyber defense. Although the NIS's National Cyber Security Center (NCSC) can operate a consolidated response system in which public and private entities can cooperate to manage and respond to crises (Regulation on Cybersecurity Duty, Article 6 bis (4)), this framework is primarily crisis-based, requiring a certain threshold to operate. In addition, the NIS and Cyber Command are entitled to engage in proactive defense separately under national law, although cybersecurity intelligence can be shared between them (Regulation on Cybersecurity Duty, Article 5). As a result, who takes the lead in proactive defense and how the authorities cooperate is not well-defined in South Korea.

Commonalities

In both countries, intelligence collection on cyber threat sources is key for successful proactive defense. **Gathering intelligence**—thereby gaining insights on adversaries' weaknesses, intentions, and capabilities—is a crucial step for defending forward as close as possible to the origin of adversary activity. In its **2024 strategy**, South Korea also highlighted an intent to bolster reconnaissance and intelligence on the sources of cyberattacks. Preemptively catching indications of malicious activities by detecting and analyzing sources is underlined as one of the subtasks of proactive cyber defense. By proactively discovering adversaries' malware and tactics and degrading their capabilities to conduct malicious cyber operations, South Korea and the United States could prevent such activities at the source before reaching their networks.

If threat hunting is limited within one nation's border, the success of proactive defense is limited as well. A proactive defense posture thus necessitates partnering with foreign countries. Indeed, the United States' Defend Forward requires the country to conduct activities **outside of U.S. networks**—both in allies' and partners' networks and in those of adversaries. Thus, the U.S. Cyber

Command views partnerships as an integral component of Defend Forward. In this vein, hunt forward operations—defensive cyber operations conducted at the invitation of a host nation—must be highlighted. Upon invitation, the U.S. Cyber Command's Cyber National Mission Force deploys a team to a host country to observe and detect malicious cyber operations there. As of March 2023, hunt forward teams had been deployed on **at least 47 missions** in more than 20 countries. With a view toward enhancing capabilities to collect and analyze threat intelligence, South Korea is **planning** to expand the exchange of threat intelligence with foreign intelligence agencies. Compared to the United States' rich experience in partnering with other countries for proactive cyber defense, including in hunt forward operations, South Korea has limited experience in this regard.

OPERATIONALIZING COOPERATION FOR PROACTIVE CYBER DEFENSE

The ROK-U.S. Alliance and Proactive Cyber Defense

At this moment, many uncertainties hang over the alliance between South Korea and the United States, as well as the paths of their respective national cybersecurity strategies. President Trump has commenced his second term in the White House and is reshaping his country's cybersecurity resources. South Korea is currently standing in the fog of uncertainty created by this political transition. Many have cast doubts on President Trump's willingness to promote U.S. alliances, including the partnership with South Korea. Furthermore, experts have been raising concerns that the absence of leadership in South Korea will put the ROK-U.S. alliance in peril.

Nevertheless, the following two principles cannot be reversed. First, the principle that the ROK-U.S. alliance applies to cyberspace must stand. Fortunately, South Korea and the United States still seem to share a common understanding that their alliance—which has lasted over 70 years—must be sustained. The foreign ministers from Seoul and Tokyo and the U.S. secretary of state met in February 2025 and discussed the necessity of enhancing the strength of their countries' alliances to ensure peace and prosperity. In addition, the United States reaffirmed its commitment to strengthening extended deterrence cooperation through **the ROK-U.S. and Japan-U.S. alliances**. In this vein, the **Strategic Cybersecurity Cooperation Framework** (SCCF), a legacy of President Yoon and President Biden, must be underscored. On April 26, 2023, the two presidents created the SCCF, signaling their agreement that the ROK-U.S. alliance applies to cyberspace. Whatever the cooperation framework, South Korea and the United States must consider this a cardinal principle.

Second, the proactive posture in cyber defense must not be abandoned. Strategic thinking evolves as cyber threats evolve; accordingly, defense has been evolving from reactive to proactive as countries adapt to the nature of cyberspace. In the words of cybersecurity experts Eric Talbot Jensen and Sean Watts, "cyberspace's **structural feature** of interconnectedness and its core condition of constant contact creates a strategic necessity to operate continuously in cyberspace." This means, as **General Nakasone** has put it, that we must not just "wait for cyber attacks to affect" our networks. In effect, it is hard to imagine that the United States and South Korea would abandon their proactive postures and reassume reactive postures, since this would go against the evolution of strategic thinking.

Laying Out South Korea's Priorities

Based on the previous analysis of the similarities and differences between the two nations' proactive approaches, this paper suggests five major considerations for South Korea to prioritize in its cooperation with the United States. Some of these are intended for both countries, while others are priorities specifically on the South Korean side.

First, the two nations need to think about a suitable cooperation framework for proactive cyber defense. Under the SCCF, the two countries undertook discussions of how the **Mutual Defense Treaty** between the United States and the Republic of Korea (MDT) applies in cyberspace. When it comes to the two countries' cooperation on proactive cyber defense, however, the MDT's role would be limited. That is because the provisions under the MDT largely address armed attack situations. The treaty contains only six articles that strengthen the two countries' efforts for collective defense. Article I outlines the two countries' commitment to the peaceful settlement of disputes and non-use of force, echoing the corresponding principles under the UN Charter. Articles II and III touch upon the two countries' cooperation when one or both are threatened by an armed attack. Article IV regulates the right to deploy U.S. forces in the territory of South Korea. The final two articles contain miscellaneous provisions on ratification, entry into force, and termination of the treaty. As illustrated earlier in this paper, proactive defense posture has been introduced in national policy mainly due to cyber threats occurring below the use-of-force threshold. Of course, proactive cyber defense is not exclusive to low-level threats, but its necessity is most pronounced at that level. Consequently, South Korea and the United States need to pursue a cooperation mechanism that enables them to proactively defend against malicious operations, including those posed below the use-of-force level.

Second, South Korea and the United States need to define the main threat actors against which they intend to proactively defend each other. According to the **2024 National Cybersecurity Strategy**, South Korea is mainly introducing an offensive response directed at threats from North Korea. Unfortunately, the real threat landscape is a bit different. For instance, pro-Russian cyber actors conducted **distributed denial-of-service (DDos) attacks** against ROK governmental websites, including that of the MND, after North Korea dispatched troops to Russia. In addition, damage caused by **China's cyber operations** has been increasingly acknowledged as severe. Whether or not to explicitly indicate threat actors in a policy document depends upon a nation's strategic calculations. South Korea and the United States, however, must identify and agree upon a common threat actor against whom proactive defense can be meaningfully employed.

To identify common threat actors, South Korea and the United States should be able to assess the feasibility and effectiveness of proactive defense measures against these threat actors, considering diverse elements like infrastructure footprints, geopolitical relations, and technical capabilities. Based on such assessments, South Korea and the United States could pinpoint shared threat actors against which collaborative proactive defense would yield a successful impact.

Third, South Korea and the United States must discuss how to successfully build collective inoculations of their networks. Proactive cyber defense aims to defend against malicious acts before they impact a country's networks; to achieve such a goal, technical capabilities must be geared

up well and outputs of technical analysis must be appropriately shared. At times, such technical measures for proactive hunting must be disseminated worldwide. To this end, cybersecurity advisories can be useful. By inoculating their own networks or global networks, the two countries can compete with malicious actors under more favorable conditions. Accordingly, South Korea and the United States must engage in dialogue on how to collaboratively raise their own technical capabilities and share these capabilities with the world for proactive defense.

Fourth, South Korea needs to rearrange its organizational framework for proactive defense and efficient intelligence sharing. Although the U.S. NSA and Cyber Command have different roles in national cybersecurity, they are directed by a single leader. This allows the United States to speedily leverage threat intelligence in cyber operations like Defend Forward. Moreover, the outputs of the command's work during hunt forward operations are appropriately **shared** with the Federal Bureau of Investigation, the Department of Homeland Security, and private companies. Releasing adversaries' malware obtained during hunt forward missions to the cybersecurity community makes that malware less effective because reactive defense can be used to detect and defeat it. In this way, the United States' proactive cyber defense can **inoculate** U.S. networks.

In South Korea, measures for proactive cyber defense are not carried out within one integrated framework. The NIS and Cyber Command are entitled to engage in proactive defense separately under national law. The consolidated response framework led by the NIS's NCSC, in which public and private entities can cooperate, functions primarily on a crisis base. This organizational framework may hinder timely intelligence sharing among relevant agencies and private entities. Considering the advantages of the United States' dual-hat leadership and its extensive experience in intelligence sharing, South Korea's organizational structure must be reexamined and coordinated.

Fifth, the legal limitations of proactive cyber defense must be explored. Proactive cyber defense entails actively detecting and analyzing the sources of malicious cyber activities. Unfortunately, these sources are physically located in other countries' territories and subject to the sovereignty of these states. Even though the international community has not reached a common understanding of how and when a state's territorial sovereignty can be breached through activities in cyberspace, South Korea needs to engage in proactive defense activities within these legal considerations. In this vein, South Korea may gain lessons from the United States' "**away game**" experiences (operations that, like Defend Forward, involve activities outside U.S. networks). The United States' experiences in away games are richer than those of other countries, and some of them are publicly available, with details of successful cases sometimes disclosed. South Korea may reference the United States' legal logic in such situations.

Conclusion

As cyber threats evolve, the tools that states devise in response evolve accordingly. Proactive cyber defense is a tool that is well-adapted to cyberspace's inherent features and threats. South Korea introduced a proactive defense posture in its 2024 National Cybersecurity Strategy, with an ultimate goal similar to the United States' Defend Forward. The concept of proactive cyber defense requires cooperation among states, as it requires tracking traces left by malicious actors globally. Given this

landscape, cooperation with the United States—the most experienced country in proactive defense—is imperative for South Korea. In conclusion, this paper, as suggested above, proposes five key areas for the two nations' cooperation to promote proactive cyber defense. South Korea and the United States need to (i) establish a suitable cooperation framework for proactive cyber defense; (ii) define the main threat actors against which they intend to proactively defend each other; and (iii) discuss how to successfully build collective inoculations of their networks. Furthermore, South Korea needs to: (i) rearrange its organizational framework for proactive defense and efficient intelligence sharing and (ii) explore the legal limitations of proactive cyber defense.

Joohui Park *is a senior researcher on the Cybersecurity Policy Research Team at the National Security Research Institute in the Republic of Korea. Donghee Kim is a senior researcher and manager on the Cybersecurity Policy Research Team at the National Security Research Institute in the Republic of Korea.*

Part III

Identifying and Assessing Tools for an Integrated Response

Active Cyber Defense in the Korean Context

By James Andrew Lewis

The Republic of Korea (ROK) faces a uniquely volatile situation in defending its networks, data, and digital infrastructure. Nuclear-armed North Korea (DPRK), unlike other leading state cyberattackers such as Russia, China, and Iran, poses a direct military threat to the ROK and makes use of missile launches, artillery fire, and (in the past) naval activity to threaten, warn, and manipulate ROK and global opinion. Drawing on one example among many, in January 2024, Kim Yo-jong, the sister of North Korean Supreme Leader Kim Jong-un, threatened an "**immediate military strike**" against South Korea in response to any "**slight provocation**." While there is a considerable degree of bluster in statements like these, the risks of taking retaliatory action against the DPRK is higher than in any other cyber conflict. This shapes any calculation of active cyber defense, defined as taking action against opponents rather than relying on attempting to deny them access to networks and data.

The international landscape for cyber defense is complicated, as all major cyberattackers are currently insulated from punitive responses—particularly from democracies, given their fear of escalation. For the ROK, however, cyber defense against the DPRK adds the risk of armed conflict—unconventional, conventional, even nuclear—to the equation. While this risk should not be exaggerated, it means that while the ROK needs a general cybersecurity strategy focused on resilience, it must also have a strategy specific to the DPRK based on active defense.

South Korea must also navigate an increasingly uncertain and hostile international environment characterized by significant instability. Expanding threats to global order include China and Russia, alongside a dangerously unaccountable DPRK. U.S. retrenchment is another factor that increases risk. This volatile landscape creates difficult challenges for South Korea's defense. Adversaries are increasingly engaging in hostile actions. While these actions have so far not involved direct military confrontation, these risks are compounded by opponent perceptions of U.S. weakness and vacillation. In response to these increasing threats, South Korea is reorienting its defense policy and strengthening alliances.

New Approaches to Cybersecurity

A critical component of this strategic evolution reflects a larger rethinking of cybersecurity. The ROK is not alone in doing this. Cybersecurity policy is being reoriented in the United States, Japan, and NATO, and concepts have evolved in ways that reflect the experience of cyber conflict. The primary change is that the emphasis on deterrence and defense is being replaced by a new focus on resilience and active defense. This is because hostile actors have not been deterred in cyberspace and defenses are too easily circumvented. New approaches to cybersecurity center on resilience—which assumes that, despite a degree of unpreventable initial success by attackers, defenders can continue providing a sufficient level of key services and recover quickly—and on active defense, which entails the use of coercive means to degrade opponent capabilities, operations, and benefits. Intelligence and espionage underpin active cyber defense.

Cybersecurity is now an integral component of national defense requirements, supported by expanded foreign intelligence gathering using cyber espionage, signals intelligence (SIGINT), human espionage, and open-source intelligence (publicly available information). While the opportunities provided by open-source collection are limited for the DPRK, it can still offer valuable insights.

Deterrence Requires Credible Threats

Cyber deterrence still has some appeal, if only because it allows governments to avoid the risk of taking action against an attacker. Deterrence reflects a defensive and passive orientation, and ultimately a desire to avoid conflict. It appealed to a cyber community more familiar with technical measures than with strategy, and to a strategic community unfamiliar with digital conflict and too oriented toward precedents from nuclear strategy. Whatever merits this defensive orientation may have had in the past, it is no longer adequate for an environment of increasing malicious cyber activity as Russia, China, North Korea, and Iran become more aggressive in cyberspace. Deterrence still makes sense as an ultimate goal, once democracies develop credible threats of punishment for malicious cyber actions, but it is inadequate until that occurs. For now, a failure to take action only encourages opponents.

The chief weakness of cyber deterrence is the lack of credible threat that would lead an attacker to recalculate benefit and risk and decide to forgo action. The DPRK is in many ways shielded from the punitive responses that democratic nations have tried so far, such as sanctions, public announcements, and law enforcement actions. The utility of sanctions in changing cyberattacker

behavior is doubtful, since the DPRK already faces an intense and encompassing sanctions regime, which its leaders shrug off as inconsequential. Sanctions and "**name-and-shame**" announcements have little effect against the DPRK. Reactive law enforcement actions can disrupt attacker networks and reduce the DPRK's gains from illicit cyber action, but do not diminish the capabilities of determined opponents. An active cyber defense policy that disrupts opponent cyber operations can provide greater protection.

Elements of Active Cyber Defense

In response to the worsening situation in cybersecurity (itself a reflection of a deteriorating international security situation), a number of countries are considering adopting or have adopted policies for active cyber defense. These responses need not be confined to cyber actions and may involve diplomatic measures, law enforcement tools, and the full range of coercive instruments available to states. Active cyber defense involves both deciding on an appropriate response for a malicious cyber action and developing adequate offensive capabilities to impose penalties. It requires more than retaliation for a specific incident, and it can use preemptive actions to prevent a cyberattack by damaging an opponent's cyber capabilities.

Active cyber defense reflects a larger adjustment by democratic nations to an increasingly hostile environment. This new environment will require significantly expanded military and intelligence capabilities, including the development of cyber offense capabilities. Active defense builds on expanded intelligence and military cyber capabilities and requires new strategies that comprise, in addition to coercive actions, coordination with allies and partners, an "all-of-government" (i.e., all pertinent ministries, not just those whose primary focus is national security) approach that combines cyber, law enforcement, financial and diplomatic tools, and designing a public narrative to accompany and explain active defense.

Active cyber defense is still ultimately defensive, as the intent is not to conquer or defeat the attacker but to reduce their ability to inflict harm and to benefit from malicious cyber action. Active cyber defense is unlikely to change an opponent's behavior (and will not deter them from malicious cyber actions), but it can disrupt their operations, degrade the returns from their cyber actions, and damage their attack capabilities. To be effective, active cyber defense cannot be limited strictly to actions in the cyber domain. It must go beyond "hacker versus hacker" and reach other key interests of the attacker, such as disrupting financial gain from hacking.

Offensive operations are only one tool among the several that are available. The fundamental elements of active cyber defense are attribution, identification of opponent targets and vulnerabilities, a menu of proportional responses, offensive cyber capabilities integrated with law enforcement and countering financial crime, a framework for collective action with other states, and the political will to act. The ROK, in adopting a policy of active cyber defense, will need to create a menu of response options derived from its capabilities that are both effective and consistent with the rules, norms, and standards governing relations between sovereign states.

Active cyber defense includes offensive cyber capabilities, which involve the ability to infiltrate another state's networks to collect information and, perhaps, to disrupt them. These actions can be preemptive, accessing and neutralizing attacker infrastructure (like command-and-control nodes or attack tools) before they are used in an attack. The ROK and its allies could consider proactive measures that involve action on the opponents' networks to disrupt attacks. While this could create the risk of retaliation or escalation of conflict, it is less risky than the increasingly dangerous cyberattacks carried out by major state opponents.

Additionally, a sophisticated diplomatic strategy is essential for engaging allies and opponents alike to manage the implications of this more active stance. Building active cyber defense on a national basis and in alliance with the United States and other partners requires the ROK to take a number of steps based on the political decision to engage in it. Active cyber defense uses the tools and levers available to states in international relations to counter transgression. Active cyber defense can be limited to those actions that do not violate sovereignty if they rely primarily on law enforcement and diplomatic measures. An initial effort may begin with these limited measures to assess risk, opponent reaction, and effectiveness and begin to build a kind of escalation ladder for responses.

EXTRATERRITORIAL ACTIONS

Extraterritoriality and sovereignty will affect decisions on active cyber defense by delineating the legal authorities required for operations. Active cyber defense could be limited to networks within the jurisdiction of the ROK, or extend to allied networks with their consent. This avoids the complication created by violating the sovereignty of the DPRK or other opponents. Active cyber defense at a minimum will require the ROK to monitor cross-border internet communications to detect signs of cyberattacks, while explicitly following privacy protections and constitutional principles.

The most difficult issue for policymakers involves preemptive cyber actions taken extraterritorially—deciding when it is justified to violate an opponent's sovereignty. Active cyber defense can require preemptive actions taken to prevent and neutralize a potential cyber threat before it can be executed by an opponent. By its nature, this means that active cyber defense can involve a violation of the opposing nation's sovereignty. Preemptive action also creates new requirements for sharing intelligence to deconflict and coordinate active cyber defense. In cases where active defense employs cyberattacks, it also creates an obligation to ensure compliance with international humanitarian law, specifically the requirements of the law of armed conflict regarding distinction and proportionality.

PROPORTIONALITY

Proportionality is one of the most important requirements of the law of armed conflict for active cyber defense. There has been discussion as to whether the proportionality of a response is determined by an individual incident or by a larger malicious cyber campaign, with the discussion moving in the direction of taking the cumulative effect of a cyberattack into account when considering proportionality. For active cyber defense to be most effective in reducing risk, it must be the latter. Active defense should not be based on tit-for-tat responses or limited to reactive actions.

Proportionality does not mean that the force used by each side must be equal. Instead, it requires balancing the necessity for effective action and the requirements of humanitarian protections. The **Geneva Conventions** define a proportional response as one that is not "excessive in relation to the concrete and direct military advantage anticipated." For active cyber defense, this includes actions that directly target those elements that contribute to an opponent's cyber capabilities while posing limited risk of collateral damage to civilian targets. This is a more forward-leaning approach, but it is necessary if active defense is to be effective. The requirement to observe proportionality does not forbid all civilian harm, and proportionality requirements can be met by better intelligence that allows for precise targeting.

Attacking DPRK critical infrastructure is probably of limited utility. The DPRK's leadership is not overly concerned with citizen well-being, and in any case, its critical infrastructure is already limited in its capacity to deliver services and poses less risk given the underdevelopment of the DPRK economy. For example, turning off power to a DPRK facility used for cyberattacks is less likely to affect civilian entities like hospitals, since they are unlikely to depend on the same power source.

However, actions against the DPRK's command-and-control hacking infrastructure and its financial networks could reduce its ability to do harm, and degrading them is a legitimate goal for active cyber defense. Identifying pressure points in the DPRK is essential for active cyber defense. Since it is unrealistic to expect to change DRPK behavior, the goal for active cyber defense should be to erode DPRK cyber capabilities and reduce their financial returns.

ATTRIBUTION

Attribution is necessary to impose consequences and perhaps ultimately create accountability. Adequate attribution is needed to ensure effectiveness, provide justification for an action to the international community, and to convince political leaders of the need for action. This is not the attribution required by a court, but what is sufficient to persuade decisionmakers that a campaign of response to malicious cyber actions is justified. It first requires identifying the nation responsible for an attack. This requires a careful calculation of what actions will be seen as both justifiable and effective. An **earlier paper** in this series discusses the requirements for attribution.

Attribution can be defined as determining the identity of an attacker. It is primarily an intelligence task. A 2022 report on UN cyberspace norms **notes that** attribution "is a complex undertaking, and a broad range of factors should be considered before establishing the source of an ICT [information and communications technology] incident." Norm 13(b) (proposed by Russia) of the 2015 report of the UN Group of Governmental Experts, the precedent for global cyber norms, **states that** "In case of ICT incidents, States should consider all relevant information, including the larger context of the event, the challenges of attribution in the ICT environment and the nature and extent of the consequences." The ability to attribute the source of an attack is critical for increasing accountability in cyberspace and giving meaning to international law and agreed norms. This concept is repeated and expanded in the 2024 **Open-ended Working Group** agreement agreed to by all member states.

The desire for a high degree of certainty in attribution before taking any action reflects concerns over the potential risks of active defense and a desire to avoid unintended consequences. The risk of escalation is overstated–there has been no **incident of escalation** in the 30-year history of cyberattacks. Escalation risk is manageable using the tools of diplomacy. The second risk is more complicated. Deciding how to act upon attribution is a political decision. Active defense is bound by both strategic implications of any action and by the ROK's commitment to international humanitarian law. Policymakers will need common understandings on the degree of certainty needed for attribution of responsibility.

The fact that the DPRK has an authoritarian mode of government does provide a benefit of sorts. First, the risk of "collateral damage" from any ROK active defense measure is lower, as DPRK assets are owned by the state. Second, the threshold for attribution of the source of an attack is much lower when the DPRK is involved. This certainty provides a greater freedom of action for the ROK and others. The North Korean state exercises very tight control over all online activities, including hacking. It can be safely assumed that no cyber action will be undertaken without its approval and involvement. There are no independent cyber actors in the DPRK or its agents deployed overseas. Once the initial question as to whether the source of an attack is North Korea is answered, more specific attribution is unnecessary for active cyber defense.

RESPONSE THRESHOLDS AND FORCE

Active cyber defense need not cross one of the most important thresholds in cybersecurity: the use of force. The emerging consensus is that an attack in cyberspace qualifies as a use of force when it produces effects equivalent to those of a kinetic armed attack, such as casualties, physical destruction, and the unacceptable disruption of data and services. There is a gray area in deciding when the disruption of services and data rises to a level of damage or destruction equivalent to the use of force, but the simplest approach is to reason by analogy and ask whether the cyber incident creates damage equal to a kinetic attack.

This definition sets a clear threshold. Most malicious actions in cyberspace have involved crime or espionage, not force. This has been a ceiling below which malicious actors in cyber conflict have been content to remain. However, as economic losses from cybercrime and cyber espionage continue to mount unabated, and as the risk of the use of cyberattacks to create harm and damage continues to grow, a forceful response by the victim state is increasingly seen as necessary.

Initially, there was speculation, such as in the 2012–2014 works of **Panetta**, **Lin**, and **Singer and Friedman**, that as states made greater use of offensive cyber operations, there could be escalation to a larger and more damaging conflict. This has not proven to be the case. In more than two decades of malicious cyber action, there has never been an incident that has led to escalation. The likely reason for this is that for attacker nations, cyber operations are part of their larger strategies of avoiding direct military conflict while pursuing their strategic objective by using unconventional means.

Instead, an implicitly observed threshold derived from the use of force determines escalation risk. No cyber action has led to casualties and only a few have led to physical destruction. Financial

harm, of which there has been a significant amount, has not triggered a forceful response. Active defense measures that do not cross the use-of-force threshold pose a much lower risk of escalation.

Escalation is managed by not crossing this use-of-force threshold, and through direct and indirect communications with the opponent. A communications strategy requires deciding what and when to tell opponents, allies, and the public, and how to tell them. The goal is to shape opponent perception and calculation of risk since it cannot be assumed that opponents infer intent from actions, and actions that escape notice have no effect.

COUNTERMEASURES AND RETORSION

In international law, countermeasures and retorsion are ways for a state to respond to the unfriendly or illegal acts of another state. An act of retorsion in international law is a measure taken by one state in response to an unfriendly or adverse act by another state. The act itself does not violate international law or customary law, even if the initial act it is responding to was either unfriendly or illegal. Retorsion thus refers to actions that are harmful to another state but are inherently lawful. Because no international laws are being broken, a state is generally free to use retorsion at any time, whether to express political displeasure or respond to a cyber incident. Examples include expelling diplomats or imposing economic sanctions. Retorsion does not require a state to prove an "**internationally wrongful act**" or meet the high legal standard needed for courts.

Countermeasures are more severe. These are actions that would normally be illegal under international law (e.g., hacking another state's networks) but become legally permissible because they are a direct response to a prior "internationally wrongful act." The response must be temporary and proportional to the injury suffered. A state cannot destroy a power grid in response to a website defacement. Countermeasures require that the target experiences pain or loss. Countermeasures (or reprisals) involve a temporary, justified breach of international law to bring an unlawful state act to an end.

Countermeasures are more likely to be effective against the DPRK than retorsive acts. Effectiveness requires identifying those places where the DPRK has something to lose, not to deter it, but to limit its ability to do further harm. This suggests that the initial targets for coercive countermeasures should be financial. Neutralizing the attacker's resources used to fund hacking operations also reduces the cyber threat. This can include seizing cryptocurrency funds tied to ransomware payments or other illegal activities. Sanctioning entities and individuals involved in the cyber campaign makes it difficult for them to transact globally. Taking down the attacker's phishing and malware delivery domains, along with supporting botnets or proxy networks used to launch cyber actions, can reduce the scale and lessen the anonymity of attacks. While the immediate goal is to impose costs for malicious cyber action, developing neutralizing capabilities can be seen as a step in developing what can be called "cyber superiority" (crippling an opponent's cyber forces and defenders in a conflict, similar to air superiority in traditional warfare) in an actual war.

Superiority in active cyber defense strategy requires degrading attacker command-and-control infrastructure used by the attacker to communicate with their malware installed on the victim's

network. This infrastructure is the nervous system for cyberattacks. Disrupting this infrastructure may produce an immediate halt to an ongoing attack and limit an attacker's ability to operate. The infrastructure can include servers (often compromised or rented) in third countries, raising extraterritoriality concerns that may require diplomatic action to address.

An attacker's assets also include the sophisticated, proprietary tools and exploits they have built. These are expensive to develop and maintain. They are also legitimate targets. Neutralizing or disrupting these capabilities inflicts a high cost. Analyzing the attacker's tools to create detection signatures and sharing the information with allies, software vendors, and service providers will force the attacker to invest significant time and money to develop new tools.

Diplomatic Aspects of Active Cyber Defense

Active cyber defense requires and can be strengthened by a robust diplomatic strategy. This involves the following:

- coordinating and ensuring deconfliction at an operational level;
- engaging with national, allied, third country, and even opponent audiences;
- developing and implementing appropriate policy, diplomatic and legal frameworks, and agreements for cooperation; and
- deciding the degree and timing of acknowledgment of active defense actions.

Multilateral coordination is absolutely essential for many active defense measures. Measures like sanctions, public announcements, and many disruptive actions can be amplified by the involvement and participation of like-minded nations. More importantly, disruptive measures need to be accompanied by steps to ensure deconfliction to reduce the risk of inadvertent interference with an operation by a friendly nation.

One benefit of the **Five Eyes** intelligence partnership is that it provides a ready mechanism for coordination among partners. Like-minded nations in Asia are not yet at the point where they could agree to a Five Eyes–style approach involving the United States, South Korea, Australia, and Japan, but existing bilateral channels can be used to coordinate and obtain support.

The disruption of cryptocurrency, money laundering, and the covert financial infrastructure that supports DPRK cyber activities is an important element of active cyber defense and requires multilateral cooperation. (And in some areas, China might be willing to cooperate.) This can be achieved by utilizing and expanding existing bilateral and multilateral efforts, such as Interpol, the Financial Action Task Force, and the **Egmont Group** of financial intelligence units.

One crucial question for policy is whether and when to inform opponents. Not informing them of responsibility for an action may be preferable in some instances, but there will also be instances where signaling an opponent serves to establish redlines and perhaps may ultimately lead to negotiations. Communicating with opponents may at first seem counterintuitive, and it can be achieved indirectly through press leaks and other mechanisms, but active cyber defense

is improved by increasing opponent uncertainty about the rewards and consequences of a cyber action.

Russia is unlikely to cooperate in any cybersecurity effort, but there may be more opportunity with China, and a diplomatic strategy should include approaching China. China is increasingly concerned with ransomware and cybercrime, and it is not always supportive of DPRK activities. China tolerates DPRK hacking activities in its territory and, at a minimum, should be asked to take action against DPRK hackers and infrastructure. While China may reject such a request or demand too high a price to accept it, putting China on notice is an essential step and could be done through expanded cooperation with allies and friends, likely Australia and Japan.

China does not appear to assist the DPRK in its malicious cyber activities, although it tolerates them. That said, China is increasingly concerned about cybercrime against Chinese companies and may be willing to discuss measures to reduce cybercrime against commercial entities. Such discussions risk, however, becoming entangled in the larger China-ROK relations and China's expectations of deference. China is a major source of espionage against the ROK. Developing a common approach to China will take time and political-level discussion that goes beyond cybersecurity.

The goal for diplomacy with allies and partners is to achieve a common situational awareness in cyberspace through information sharing, enabling partner nations act on threats. Building shared situational awareness is fundamentally a political challenge. Sharing data may require formal agreements to ensure compliance with ROK law. Both the ROK and the United States will benefit if they can agree on a **Cyberattack Severity Classification Framework** and joint rules of engagement to make coordination easier. The ROK may also benefit from creating liability protections for private sector data sharing similar to the U.S. Cybersecurity Information Sharing Act.

The degree of government cooperation and the timing of coordination efforts also require political-level decisions: Are allies and partners informed of an action in advance, at its onset, or afterward? Existing intelligence, military, law enforcement, and diplomatic channels can be used to facilitate this coordination. Prior agreement on consultations (building on existing agreements) is better than ad hoc notifications and can be built as additions to existing coordination mechanisms.

ALLIES AND PARTNERS

Active cyber defense can be seen as part of a larger shift in defense policy. It treats cyber operations as part of a more robust and comprehensive defense strategy. This proactive approach allows for the preemption of attacks, though this will require the careful integration of new technologies and strict management of escalation risks. To be most effective, this strategy relies heavily on coordination with allies to maintain a common approach to addressing cyberattacks.

The U.S.-ROK alliance has made cybersecurity and information sharing a core pillar of cooperation. Overall, the effectiveness of existing cooperation mechanisms is improving, driven by shared threat

perceptions and high-level commitments. Cooperation is based on a tiered structure, from the strategic to the operational.

Coordinated attribution of malicious activity will require better information sharing between partners and, perhaps, new mechanisms for sharing and harmonization of policies. Additionally, creating a framework of technical and factual attribution would help frame the development of active cyber defense actions intended to disrupt opponent cyberattack capabilities. Even with this, however, attribution and proportionality remain sovereign decisions.

INTELLIGENCE SHARING

Active cyber defense blends law enforcement, military, and intelligence activities. Intelligence coordination is essential, if only because the ROK's major opponents rely heavily on their intelligence services—Russia's Main Intelligence Directorate (still referred to as the GRU) and Federal Security Service (FSB), China's Ministry of State Security and People's Liberation Army, and Iran's Islamic Revolutionary Guard Corps—to undertake hostile cyber actions. Decisions on appropriate responses to cyberattacks raise the issue of how to incorporate offensive cyber into national and alliance military doctrine. One such issue for active cyber defense comes from the differing capabilities of the partners. Another is the dilemma of sharing sensitive information. Finally, as discussed later, some active cyber activities may be covert. While the goals of a larger diplomatic strategy can be established, the implementation of that strategy will need to be dynamic and evolve with opponents' reactions.

Law enforcement actions may have greater effect if they target the DPRK's financial operations. The primary incentive for DPRK hacking is financial gain, especially through cryptocurrency, given its anonymity and ease of transfer. North Korea uses cyber-enabled theft, money laundering, extortion campaigns, and crypto theft to fund projects. Crypto thefts often occur from exchanges and wallets where users keep their coins, using a combination of tactics including phishing and malware to gain access. Interdiction (when possible) of the DPRK's fraudulent IT workers and those who enable their remote activities would also damage North Korea's cyber operations.

This makes active cyber defense primarily an intelligence battle. The DPRK's hacking initiatives are orchestrated by the Reconnaissance General Bureau (recently expanded and renamed the Reconnaissance Information General Bureau), North Korea's intelligence agency, with **reportedly** at least 6,000 cyber operatives (organized as "Bureau 121") carrying out operations against banks and cryptocurrency exchanges in at least 17 countries. North Korea funnels funds from cybercrimes to its weapons of mass destruction programs. **Cryptocurrency exchanges** are a favored target for North Korea, **since they allow the DPRK** "to generate income in ways that are harder to trace and subject to less government oversight and regulation than the traditional banking sector." Ransomware is another preferred DPRK tool. The use of cryptocurrencies and ransomware has become a central element of DPRK cybercrime operations.

Recommended Policy and Operational Measures

The requirements for a new and more effective approach to cybersecurity for the ROK can be summarized as follows:

Legal and Policy Framework

- Develop or revise legal authorities and policies, consistent with international humanitarian law and emphasizing retorsion and countermeasures.
- Define legal authorities required for extraterritorial operations.
- Establish a framework for determining responses, based on the cumulative effect of cyber campaigns rather than isolated incidents.

Intelligence and Attribution

- Expand national intelligence capabilities for cyber espionage using SIGINT, imagery, open-source intelligence, and human intelligence (HUMINT).
- Define the parameters of the "degree of certainty" required for political leaders to authorize action (and avoid court-style evidentiary requirements).
- Map the DPRK's "pressure points," including command-and-control nodes and financial networks.
- Create a shared intelligence pool with the United States and other allies to deconflict ongoing operations.

Operational Capabilities

- Develop and maintain tools to infiltrate and neutralize attacker infrastructure before attacks are launched (preemption). The U.S. Cyber Command's 2018 Joint Task Force ARES is a precedent.
- Expand capabilities to seize cryptocurrency and disrupt illicit money laundering networks.
- Establish mechanisms to share with private sector partners and create legal safeguards for the private sector to share data with the government.

Diplomatic Coordination

- Coordinate with allies and regional partners (e.g., Australia and Japan) for "name-and-shame" campaigns, economic penalties, and, in some cases, offensive operations.
- Strengthen ROK-U.S. joint rules of engagement and create a severity classification framework.
- Design a strategic communication plan to explain and justify active defense measures to the global community.

Conclusion

Successfully implementing the requirements for active cyber defense depends on sustained political will to take new risks. The ROK can set a precedent for future security policy by endorsing active cyber defense and, if implemented effectively, it can reduce losses from DPRK cyber actions. Active cyber defense is not an all-or-nothing proposition, and the legal and policy foundations can be laid and actual implementation carried out on an incremental basis to accommodate political perception of risk.

Will this stop the DPRK or other malicious cyber actors? Not at all. The profits from hacking are too important for Pyongyang to give them up. Russia, the DPRK, and China are hostile to democracies. But active cyber defense can reduce the harm and cost of malicious cyber action.

James A. Lewis *is a senior adviser (non-resident) in the Economic Security and Technology Department at the Center for Strategic and International Studies in Washington, D.C.*

South Korea's Integrated Cyber Defense Framework

Active Cyber Defense and Reactive Responses

By Sunha Bae

Introduction

Cyberspace is structurally favorable to attackers. There is little cost when they fail and significant gains when they succeed. As a result, traditional deterrence is difficult to achieve in cyberspace, and the imbalance between the risks and rewards of cyberattacks persists. For this reason, **international policy** and **research** have emphasized approaches that make an attacker's calculus less advantageous, generally though strategies that reduce the cost-effectiveness of conducting cyber operations.

The structural advantage that cyberspace provides to attackers affects many states. North Korea, in particular, uses cyber warfare as a core component of its military strategy, persistently engaging in large-scale cryptocurrency theft, exfiltration of defense technology and state secrets, and attempts to disrupt critical infrastructure. In the first half of 2025 alone, North Korean hackers allegedly stole approximately $1.6 billion in cryptocurrency—around **70 percent** of global cryptocurrency theft during that period. Moreover, illicit cyber activities now account for roughly **30 percent** of North Korea's foreign currency earnings and, aside from weapons sales to Russia, constitute its primary source of external revenue—funds that directly support its military buildup.

For South Korea, North Korea's cyber activities represent a core threat to national security and economic stability, extending beyond espionage and illicit financial gain. At the same time, **China** conducts extensive cyber espionage campaigns that target strategic industries such as semiconductors, batteries, and telecommunications, while **Russia** has expanded cyber activities against South Korea alongside deepening political, military, and economic cooperation with North Korea. Emerging threats—including ransomware operations, supply chain intrusions, and AI-driven automated attacks—further underscore the limits of South Korea's traditional posture of passive defense.

Although the South Korean government has repeatedly attributed and condemned North Korea's cyber activities, meaningful cost-imposing responses remain limited. Attributing attacks without follow-on measures raises doubts about the state's capacity and erodes the credibility of its deterrent posture. As former U.S. National Intelligence Officer for Cyber Issues Sean Kanuck **noted** during a visit to Singapore in 2016, once the U.S. government formally accused Russia of interfering in the election, it needed to act. Inaction after attribution creates a negative precedent—undermining both retaliatory threats and the credibility of cyber capabilities. The United States responded with diplomatic expulsions, **economic sanctions** on Russian intelligence agencies and affiliated individuals and entities, and, later, disruption of the Internet Research Agency (IRA) network ahead of the United States' 2018 midterm elections. Although these measures did not fully deter or halt Russia's election interference efforts, they helped prevent doubts about the integrity of the 2018 midterm results and heightened domestic and international awareness of election security. Moreover, they contributed to a broader trend in which democracies such as the United Kingdom, Canada, and Taiwan began publicly disclosing and responding to foreign interference in their own elections.

Repeatedly attributing a wide range of cyber incidents primarily to North Korea raises questions about analytical objectivity, particularly in cases where the involvement of other state or non-state actors cannot be entirely excluded. In such circumstances, attribution risks becoming a declaratory act rather than a credible instrument of statecraft. Without tangible follow-on measures, it may instead signal tolerance for continued malicious activity. South Korea therefore needs a more comprehensive and integrated defense framework to address recurring cyber threats. In this context, the concept of active cyber defense (ACD) becomes relevant.

A 2025 CSIS white paper, "**Forging Forward: South Korea's Proactive Cyber Defense and Strategic Cooperation with the United States**," focused on proactive cyber defense, emphasizing the need to detect, disrupt, and block threats before attacks occur. That paper highlighted the limitations of South Korea's traditionally passive posture and underscored the need for a more forward-leaning cyber strategy.

Building on that foundation, this study puts forth an Integrated Cyber Defense Framework that combines ACD with reactive response measures. ACD refers to proactive disruption and persistent engagement as a mode of direct cyber intervention, while reactive response measures encompass sanctions, law enforcement, diplomacy, and technical actions. The proposed Integrated Cyber Defense Framework aims to establish a defense model tailored to South Korea's strategic

environment and recommend improvements in institutional foundations, governance, and cooperative structures that will be necessary to implement such an approach.

Conceptualizing South Korea's Integrated Cyber Defense Framework

ACD RESEARCH TRENDS AND CHARACTERISTICS

Before defining the Integrated Cyber Defense Framework, it is necessary to examine how ACD has been defined in existing research and policy discourse. As the most conceptually contested component of the framework, ACD warrants particular attention.

ACD originated from the military concept of active defense and has been subsequently adapted to the cyber domain, with the United States playing a leading role in its early conceptualization at both the military and policy levels. In the 2011 **Strategy for Operating in Cyberspace**, the U.S. Department of Defense (DOD) formally defined ACD as "synchronized, real-time capability to discover, detect, analyze, and mitigate threats and vulnerabilities." This definition framed ACD as an operational capability distinct from static or purely preventive defense, and it was later reflected in **U.S. Cyberspace Operations Doctrine**, in which ACD was explicitly distinguished from passive cyber defense.

Despite this early institutionalization in U.S. doctrine, the scope of ACD has remained contested. Some scholars define ACD as a **counterpart to passive defense**, encompassing both internal mitigation and actions conducted beyond the defender's own network. Others focus instead on the direct and persistent mode of **intervention**, emphasizing deliberate and direct defensive actions that **disrupt or constrain** adversary operations and shape behavior below the threshold of armed conflict. Another line of analysis situates **ACD in the gray zone** between passive defense and offensive cyber operations. More expansive approaches conceptualize ACD as an overarching framework integrating **reactive, heuristic, and proactive modes** of cyber defense. As a result, ACD lacks a stable and universally accepted definition.

Nevertheless, across these varying interpretations, a common feature emerges: ACD entails an intervention-oriented posture that seeks to dismantle or disrupt threats, rather than merely aiming to strengthen defensive measures or allow attacks to be endured. In this sense, the distinctive feature of ACD lies in its emphasis on deliberate actions that directly shape or constrain an adversary's operations, distinguishing it from traditional passive defense.

DEFINING AN INTEGRATED CYBER DEFENSE FRAMEWORK FOR SOUTH KOREA

Building on this conceptual clarification, this report proposes an Integrated Cyber Defense Framework to structure South Korea's national cyber response around two complementary components: ACD and reactive response. The distinction between these components is not based simply on whether an attack has occurred, but on differences in purpose and mode of intervention. All activities within this proposed framework remain firmly within the domain of defense, not offense.

- **Active cyber defense** aims to identify, disrupt, and dismantle threats. ACD includes internal measures such as threat hunting, vulnerability mitigation, and honeypot operations, but it also includes external measures such as the disruption of adversary infrastructure and the use of persistent engagement or international coordination to detect and constrain malicious activities.
- **Reactive response** seeks to minimize damage after an incident, strengthen resilience, and impose consequences on the attacker. It includes detection, recovery, resilience enhancement, sanctions, law enforcement actions, diplomacy, and technical measures to impose costs.

The objective of this framework is to establish a national defense posture in which threats are actively **constrained and disrupted** while **losses are minimized and consequences are imposed on malicious actors**. Within this context, the term "offensive," used in South Korea's 2024 **National Cybersecurity Strategy**, is prone to misunderstanding and may be read to imply operations that exceed defensive intent. Likewise, "proactive" alone does not adequately capture the full spectrum of measures needed. Accordingly, within this framework, ACD is defined in a focused sense to reduce conceptual ambiguity and clarify its role as a defensive function focused on disrupting or constraining adversary operations.

Trends in Cyber Defense Among Major Countries

UNITED STATES

Strategy and Governance

The United States has established a strategy that combines ACD and reactive response measures for cyber defense. This approach is guided by the objective of disrupting attacks before they occur, imposing costs after attacks, and recovering quickly.

The DOD defines **ACD** as a form of cyber defense that actively detects and blocks malicious activity before it affects U.S. networks and systems. These activities include **disrupting** or dismantling malicious cyber activities at their source, including actions conducted below the threshold of armed conflict. In other words, U.S. policy considers certain offensive cyber operations as part of active defense when their purpose is to preemptively defend against or weaken attacks targeting U.S. interests.

Since 2018, this ACD approach has evolved into a strategy of **persistent engagement,** which involves continuously countering adversaries across the cyber domain and, when necessary, operating in foreign networks to disrupt or block adversary infrastructure through **defend forward** and **hunt forward** operations. The 2020 **Cyberspace Solarium Commission** emphasized deterrence by denial through improved defensive capabilities, as well as cost imposition through continuous responses using all instruments of national power, recommending ACD as a core component of this approach. While the 2023 **U.S. National Cybersecurity Strategy** reaffirmed this approach, the Trump administration's 2026 **Cyber Strategy for America** places greater emphasis

on imposing costs on adversaries and adopting a more assertive posture toward malicious cyber actors, while emphasizing the active use of both defensive and offensive cyber operations.

The United States' legal and institutional framework supports this approach. The **National Defense Authorization Act** (NDAA) empowers the DOD to take "appropriate and proportional actions" in foreign cyberspace against "active, systematic and ongoing campaigns of attacks" conducted by key adversaries such as Russia, China, Iran, and North Korea. In addition, **National Security Presidential Memorandum 13** (NSPM-13), published in 2018, although not publicly released, is reported to have streamlined decisionmaking for time-sensitive offensive cyber operations, allowing the DOD to respond more quickly to malicious cyber activity. However, significant or sensitive cyber operations still require **coordination** with the White House and relevant departments. The **FY 2019 NDAA** also requires quarterly congressional reporting on the use of cyber weapons and cyber operations, in order to enhance transparency and strengthen congressional oversight.

The organization responsible for planning and conducting ACD operations is **U.S. Cyber Command (USCYBERCOM),** within the DOD. USCYBERCOM integrates personnel from military services and intelligence agencies to build cyber defense and operational capabilities, organized into cyber operations, cyber defense, and research and development. Moreover, the United States is currently considering establishing a separate **Cyber Force** to ensure that USCYBERCOM secures sufficient professional personnel and resources to execute long-term cyber strategy. These discussions also encompass the need to cultivate and secure cyber-specialized **legal experts** who can support the planning and execution of complex cyber operations.

In terms of reactive response, resilience enhancement and cost imposition are key. Legally, the Federal Bureau of Investigation (FBI) and the Department of Justice (DOJ) criminally indict malicious actors and publicly disclose their identities to limit their operational freedom. Economically, the Department of the Treasury (USDT) freezes the assets of international hacking groups and their enablers and blocks their access to the global financial system. Diplomatically, the Department of State (DOS) and the White House attribute cyberattacks to specific states together with allies and issues joint attribution and public condemnations, linking these with diplomatic sanctions to degrade relations with the attacking state and increase the perpetrators' international isolation. The Cybersecurity and Infrastructure Security Agency (CISA) and the National Security Agency (NSA) conduct incident response, enhance resilience, protect critical infrastructure, and issue security advisories. Through this integrated approach, the United States brings together technical, law-enforcement, diplomatic, and economic measures to impose meaningful costs on malicious actors.

U.S. Cases

The United States is currently the most active country in conducting both active and reactive cyber responses. Beginning in 2018, it has increasingly publicized ACD operations, strategically signaling its tracking capability and willingness to respond. International cooperation with the European Union, **NATO**, and Five Eyes partners has also steadily expanded.

The United States has utilized a range of tools as a part of its response, such as disrupting infrastructure, taking down hacking groups, attributing attacks publicly, issuing security advisories, enacting sanctions, and issuing indictments of internal and external collaborators. A prominent example is the **SolarWinds incident** in 2020, when the FBI, NSA, CISA, and the Office of the Director of National Intelligence (ODNI) jointly attributed the attack to the Russian government, followed by diplomatic expulsions by the DOS, sanctions by the USDT, and insider trading indictments by the Securities and Exchange Commission (SEC). In the 2024 **Volt Typhoon** case involving China-linked actors, the FBI, NSA, and CISA jointly identified the threat and issued a security advisory, while the DOJ indicted the hackers and the USDT sanctioned associated companies and individuals. The FBI and CISA also carried out proactive technical measures to identify and block the infrastructure used in the attack. The United States has also undertaken coordinated international operations against non-state actors such as the **Hive ransomware group** in 2023, resulting in global tracking, system takedowns, and arrests.

Table 1: U.S. Cyber Response Cases

Year	Entities	Response	Type
2016	**White House**, ODNI, DOS, USDT, DHS, FBI, **CYBERCOM**, NSA	Sanctions on related entities and individuals; expulsion of diplomats; joint attribution and condemnation; disruption of internet access for the IRA, Russia's disinformation organization	Active + reactive
2020	**USDT**, DOJ, FBI, **CYBERCOM**, CISA	Temporary disruption of the Russia-based major botnet Trickbot; joint attribution and condemnation; sanctions and indictments	Active
2020	**White House**, ODNI, DOS, **USDT**, DOJ, FBI, NSA, CISA, **SEC**	In response to the SolarWinds supply-chain attack: Senior political intervention through a presidential-level call urging specific actions; official attribution to Russia's SVR, Russia's civilian intelligence agency; expulsion of Russian diplomats; prohibition of Russian sovereign bond issuance in U.S. financial markets; sanctions on related organizations and companies; coordination with the United Kingdom and others	Reactive
2022	**CYBERCOM**	Deployment of a hunt forward team to Ukraine; pre-incident detection and removal of malicious activity on Ukrainian networks	Active
2022	**DOJ**, FBI, **CISA**, NSA	Court-authorized disruption of the Russian GRU-controlled global botnet Cyclops Blink; joint attribution; security advisory; Five Eyes cooperation	Active

2022	DOS, **USDT**, FBI, **CYBERCOM**, **CISA**	In response to cyberattacks on Albanian government agencies: public attribution to Iran's Ministry of Intelligence and Security; security advisory; sanctions; indictments	Reactive
2023	**USDT**, **DOJ**, FBI, CISA	Infiltration of the Hive ransomware group; acquisition of decryption keys; international coordination to seize and shut down Hive servers and websites; sanctions; joint attribution; security advisory; Five Eyes cooperation	Active
2023	**DOJ**, FBI, CYBERCOM, CISA, NSA	International operation to disrupt and eliminate the Russia-based cyber espionage malware, Snake; joint attribution; security advisory; Five Eyes cooperation	Active + reactive
2023	**DOJ**, FBI, CISA	International disruption of Qakbot malware; seizure of approximately $8.6 million in stolen cryptocurrency; indictments	Active + reactive
2024	**DOJ**, **CISA**, NSA	In response to cyberattacks involving the Microsoft Exchange server exploitation: joint attribution; infrastructure takedown; international coordination	Active + reactive
2024	ONCD, **DOJ**, **USDT**, **FBI**, CYBERCOM, **CISA**, NSA	In response to the Volt Typhoon attack: joint attribution; security advisory; disruption of the Chinese botnet used in the attack; indictments and sanctions against APT 31; Five Eyes cooperation	Active + reactive

Source: Author's analysis.

UNITED KINGDOM

Strategy and Governance

The United Kingdom uses the term ACD in a more limited sense than the United States. The United Kingdom has operated an automated defense system, called the **Active Cyber Defence** (ACD) program, since 2016, which is designed not to neutralize adversaries but to prevent and reduce harm from their cyberattacks and strengthen collective defense across the digital ecosystem. The National Cyber Security Centre (NCSC) states that the ACD program aims to contain damage quickly even when serious attacks occur and improve resilience against routine threats.

However, despite differences in terminology, the United Kingdom's cyber response strategy also combines elements of ACD and reactive measures. The strategy includes the ACD program as well as offensive cyber operations. In 2020, the United Kingdom established the **National Cyber Force** (NCF)—a joint organization of Government Communications Headquarters (GCHQ) and the Ministry of Defence (MOD)—to institutionalize cyber operational capabilities. Unlike the NCSC,

which provides and leads domestic cyber resilience, the NCF conducts offensive cyber operations against a wide range of targets, including terrorist groups, hostile states, and criminal organizations. In addition, the United Kingdom announced plans in 2025 to establish a separate **Cyber and Electromagnetic (CyberEM) Command**, which handles defensive, cyber, and electronic warfare missions across military and government domains so that the NCF can focus more specifically on offensive and tactical cyber operations.

However, the United Kingdom does not officially classify these activities as part of ACD. Instead, it designates them as offensive cyber operations or counter-cyber operations. While using the term "offensive," the United Kingdom emphasizes transparency and responsible behavior in cyberspace in accordance with domestic and international law. Although the NCF's operations are conducted covertly, the United Kingdom has publicly explained how the NCF conducts cyber operations to signal clearly that the country has both the capability and the willingness to impose costs on adversaries.

Because the NCF is a joint organization between GCHQ and the MOD, it is subject to both the **Intelligence Services Act** (ISA) and defense-related laws and is overseen by political, parliamentary, and judicial mechanisms. Politically, the United Kingdom established ministerial oversight through the ISA, which legally defines the mission of intelligence agencies. The NCF also requires the approval and direction of the prime minister and relevant ministers, reflecting the principle of ministerial responsibility, which prevents the military and intelligence agencies from independently abusing the power to conduct cyberattacks.

Second, parliamentary oversight is exercised through the Intelligence and Security Committee (ISC), which oversees intelligence agencies. Similarly, the NCF reports its activities to the ISC, enabling Parliament to check executive authority and ensure democratic accountability.

Third, judicial oversight is applied to NCF intelligence operations that require investigative powers, which fall under the authority of the **Investigatory Powers Act** (IPA) and the supervision of the Investigatory Powers Commissioner (IPC). The IPA regulates interception, equipment interference, and bulk data collection; such activities require both ministerial authorization and IPC approval. The United Kingdom also stresses responsible, proportionate, and voluntary adherence to international norms in **NCF** operations, reinforcing the legitimacy of its cyber military and intelligence activities domestically and internationally.

In terms of reactive response, the National Crime Agency (NCA) conducts investigations into and prosecutions of cybercriminals through international cooperation. However, the United Kingdom generally prefers practical **sanctions** over symbolic indictments of foreign-based hackers, as the Crown Prosecution Service (CPS) has high evidentiary requirements and the likelihood of arresting foreign cybercriminals is low. To support sanctions, the United Kingdom established a standalone **cyber sanctions regime** in 2020, designating cyberattacks as a separate sanctionable category. This framework allows sanctions not only on individuals residing in the United Kingdom but also on external enablers and service providers, signaling an intention to affect overseas networks.

The **Foreign, Commonwealth & Development Office** (FCDO) attributes cyberattacks to specific states and issues public statements of condemnation, imposing diplomatic costs on responsible states. In recent years, the United Kingdom has actively coordinated attribution and sanctions with the United States. Joint sanctions by the two countries signal to the international community that cybercrime is a global issue and demonstrate that such attacks are traceable, thereby increasing the deterrent effect on malicious actors.

UK Cases

The United Kingdom possesses offensive cyber capabilities and publicly acknowledges their existence, but publicly released examples of offensive cyber operations are extremely limited. The ACD program includes proactively blocking attacks, removing malicious websites in real time, and shutting down harmful infrastructure, referred to as the **Takedown Service**. Between September 2024 and August 2025, the UK government **reports** removing 12,000 cyber campaigns, stopping 26,000 phishing campaigns, detecting and resolving 79 percent of phishing attacks within 24 hours, and mitigating 50 percent of such attacks within one hour. As of August 2025, **sanctions** had been imposed on approximately 70 individuals and nine organizations, in addition to the cases listed in Table 2.

Table 2: U.K. Cyber Response Cases

Year	Entities	Response	Type
2017	**GCHQ**, **MOD**	Cyber operations against ISIS; disruption of the terrorist organization's online activities, equipment, and networks	Active
2023	**FCDO**	Sanctions on 11 Trickbot/Conti ransomware members; asset freezes and travel bans; indictments by the U.S. DOJ	Reactive
2023	**FCDO**, NCSC, NCA	Public condemnation of Russian cyber interference and hacking targeting political and democratic processes; summoning the Russian ambassador; indictments of related individuals	Reactive
2023	**NCA**	Arrested more than 120 individuals affiliated with the Genesis Market operation; international cooperation with the United States and the Netherlands	Reactive
2025	FCDO, **NCSC**	Sanctions on GRU units—Russia's military intelligence agency—and 18 Russian individuals for continued malicious cyber campaigns	Reactive
2025	**NCSC**	Public condemnation of cyber espionage activities targeting critical infrastructure by a Chinese hacking group	Reactive

Source: Author's analysis.

JAPAN

Strategy and Governance

Japan formally introduced the concept of ACD through a **legal amendment** in May 2025, with full implementation scheduled for 2027. The **law** defines ACD as measures that prevent cyberattacks threatening national security and limit their spread by collecting and identifying information prior to an attack and blocking the attacker. It further authorizes administrative interception for ACD purposes, focusing on cross-border data collection based on the recognition that most cyberattacks originate outside Japan. The law also provides for establishment of an **independent oversight body** to review compliance with privacy and communications secrecy protections.

Alongside these legal and institutional changes, Japan has been reorganizing its relevant organizations to support the transition toward ACD. In July 2025, the National Center of Incident Readiness and Strategy for Cybersecurity (NISC) was expanded and reorganized into the **National Cybersecurity Office** (NCO), which was designated as the control center responsible for coordinating and implementing Japan's shift to ACD. In the December 2025 **Cybersecurity Strategy**, Japan referred to ACD for the first time in a strategic document and emphasized the need to impose costs on malicious actors.

Under the NCO's direction and coordination, the Japan Self-Defense Forces (JSDF) and the National Police Agency (NPA) **serve** as operational units, while the Ministry of Foreign Affairs (MOFA) must be consulted in advance when measures involve foreign systems. Approval procedures have been established to ensure that such measures do not exceed what is necessary and are conducted under fair and appropriate processes. In the event of a sophisticated cyberattack on the government or critical infrastructure, the prime minister can designate the incident as a "specified incident" and authorize the JSDF to respond. Outside of emergencies, the JSDF must obtain prior **approval** from the independent telecommunications information oversight commission. But the law does not explicitly clarify whether such operations can be conducted from overseas locations, and Japan has framed ACD as a defensive—not offensive—measure to avoid controversy over potential violations of international law.

Japan is also strengthening the JSDF's capabilities for ACD. The JSDF Cyber Defense Command (JCDC) was established in March 2022 as a direct reporting unit under the Minister of Defense (MOD), integrating cyber personnel from all three JSDF branches. Japan **plans** to expand the specialized JCDC to 4,000 personnel by 2027—up from roughly 530 in 2022. Including information technology (IT) personnel involved in system acquisition and maintenance, the total cyber-related workforce of the JSDF is projected to reach approximately 20,000 personnel.

However, because the law has not yet entered into full effect, there are limited publicly available cases that demonstrate the Japanese government's operational direction. In addition, Japan's **Defense Capability Build-Up Plan**—one of the country's three major national security strategies—focuses primarily on cybersecurity within the MOD, JSDF, and defense industries; early detection and response to threats; and the development of capabilities to disrupt adversary use of cyberspace. In comparison to the United States and United Kingdom, Japan's emphasis is somewhat limited. **Constitutional concerns** regarding privacy and communication secrecy have also led the

May 2025 law to impose strong restrictions and independent oversight on ACD measures, which may make the implementation of operational measures more difficult.

In terms of reactive response, Japan has pursued diplomatic measures such as financial sanctions, joint attribution, and publication of security advisories, as well as law enforcement cooperation through joint investigations and indictments. Criminal investigations are primarily led by the NPA, with additional support from the Japan Cybercrime Control Center (JC3). Japan currently maintains a complete **ban on trade** with North Korea and has imposed asset **freezes** on individuals and entities, including hacker groups. Under the **Foreign Exchange and Foreign Trade Act**, ransom payments to sanctioned actors are also prohibited. Financial sanctions are primarily administered by the Ministry of Finance (MOF).

However, toward China, Japan has generally favored joint attribution, security advisories, and diplomatic warnings based on international cooperation, rather than direct sanctions. Although there is no publicly disclosed national procedure for public attribution, the MOFA has broadly led these efforts and has expressed its intention to expand cooperative and national-level attribution activities.

Japan Cases

Japan has relied more heavily on reactive responses such as public attribution, indictments, and sanctions. According to a 2023 report by the German think tank Interface, Japan has used public attribution in approximately **six cases** since 2014, primarily using political and technical attribution. While Japan has not taken a leading role in global joint investigations or ACD operations, recent cases show that it has cooperated with the United States and European countries in joint investigations of cybercriminal organizations and infrastructure takedown operations.

Table 3: Japan Cyber Response Cases

Year	Entities	Response	Type
2024	NPA, **MOF**	Public attribution of the theft of approximately ¥48 billion in Bitcoin from the Japanese crypto exchange DMM, identifying the North Korean hacking group Lazarus; calling for G7 cooperation	Reactive
2021	**NPA**	Public attribution of intrusions into Japan Aerospace Exploration Agency (JAXA) and defense industry firms, identifying a Chinese military-linked hacking group; indictment of a Chinese national	Reactive
2023	**MOFA**	Public attribution and sanctions on three North Korean hacking groups and four affiliated individuals by freezing assets	Reactive

2025	NPA, **NCO**	Public attribution of cyberattacks targeting advanced technology and security-related institutions, identifying the China-linked hacking group MirrorFace	Reactive
2025	**JC3**	Public attribution and multinational cooperation with the United States and European partners to dismantle the infrastructure of the Lumma Stealer malware; joint investigation; coordinated infrastructure takedown	Active + reactive

Source: Author's analysis.

AUSTRALIA

Strategy and Governance

Australia is a country that, like the United Kingdom, openly acknowledges and conducts offensive cyber operations. These operations are carried out by the **Australian Signals Directorate** (ASD), the intelligence agency responsible for cybersecurity, offensive cyber operations, and intelligence collection. The ASD has publicly stated that it has these **capabilities**.

Under the **Intelligence Services Act 2001**, which defines the mandate of intelligence agencies, the ASD holds the authority to prevent and disrupt foreign cybercrime. Based on this legal foundation, the ASD conducts operations that disrupt the infrastructure of foreign cybercriminals, including intrusions targeting overseas networks. Although government authorization is required to conduct such operations, the specific orders and approval procedures vary depending on whether the offensive cyber capability is being used to support military operations, assist law enforcement, or deter and respond to cyber threats. These operations are also subject to the ASD's existing legislative and oversight framework, including independent oversight by Australia's **inspector-general of intelligence and security**, thereby ensuring legal legitimacy and accountability. Australia has also announced ongoing **joint operational plans** between the ASD and Australian Federal Police (AFP) for offensive cyber operations, and has emphasized that the Australian government will counterattack criminals and hackers when cyber incidents occur.

Australia is advancing the **Resilience, Effects, Defence, Space, Intelligence, Cyber, Enablers** (REDSPICE) program, which aims to expand and strengthen the ASD's intelligence, offensive, and defensive cyber capabilities. The government plans to invest approximately $10 billion over 10 years. REDSPICE includes plans to triple Australia's current offensive cyber capability; double continuous cyber hunting activities; employ advanced AI, machine learning, and cloud technologies; and expand global operational reach fourfold.

The ASD originated as a defense signals agency and, broadly speaking, functions as an independent entity within the Australian government's Department of Defence. It is not part of the Australian Defence Force (ADF) chain of command, but **reports directly** to the Department of Defence and supports ADF operations. Within the ADF, there are several cyber-related units, including a Cyber Command responsible for operational cyber activities, a **Cyber Operations Division** responsible for the integrated management and security of military cyber and communications infrastructure,

and a Space and Cyber Capabilities Division responsible for capability development and acquisition. This close and clearly defined relationship between the ASD and Department of Defence provides a coherent governance structure for conducting offensive cyber operations.

Investigation and judicial responses are led by the **Australian Federal Police** (AFP), which conducts cybercrime investigations, cooperates with state police and international law enforcement, and carries out prosecutions. Sanctions are designated and announced by the Department of Foreign Affairs and Trade (DFAT) with the approval of the attorney general. In the 2025 **Australian Cyber Response Plan** (AUSCYBERPLAN), law enforcement and attribution were designated as official government response activities. Additionally, in 2021, Australia established a significant cyber incident **sanctions** framework, enabling direct sanctions—asset freezes, prohibitions on providing economic resources, and travel bans—against actors responsible for such incidents.

DFAT leads the **Cyber Rapid Assistance for Pacific Incidents and Disasters** (RAPID) team, which can be immediately deployed to Pacific partner countries that experience cyber incidents to assist with recovery and response, thereby extending Australia's cyber response capability beyond its borders.

Australia Cases

Australia is one of the few countries that officially acknowledges possessing offensive cyber capabilities, though publicly disclosed cases remain limited. Since the partial disclosure of the ASD's offensive cyber capabilities in support of military operations in **Iraq and Syria in 2016**, most offensive cyber operations have been conducted covertly. However, as a member of the Five Eyes alliance—alongside Canada, New Zealand, the United Kingdom, and the United States—Australia actively engages in international joint investigations and ACD activities.

Although Australia has generally taken a more cautious approach to public attribution than the United States and the United Kingdom, it adopted a more active posture in 2024. DFAT and the ASD's Australian Cyber Security Centre (ACSC) jointly led international cooperation efforts to attribute and issue **security advisories** regarding China's cyber espionage activities. In 2024, DFAT officially designated China as a cyber threat actor, leading to diplomatic friction with the **Chinese embassy**.

Table 4: Australia Cyber Response Cases

Year	Entities	Response	Type
2018	**DFAT**	Joint attribution and condemnation of the cyber espionage activities of China's state-sponsored hacking group APT 10	Reactive
2020	**Department of Defence**, ASD	Disruption and dismantling of cybercriminal infrastructure during the Covid-19 pandemic	Active

2022	**AFP**	Public attribution of a Russia-based cybercriminal group's responsibility for the Medibank data breach; sanctions	Reactive
2023	**AFP**	Disruption of servers of the Russia-linked ransomware group BlackCat; acquisition of decryption keys; international coordination with the United States and Europe	Active
2024	**DFAT**, **AFP**	Sanctions (asset freeze and travel ban) on Russian leaders of the ransomware group LockBit; jointly conducted Operation Cronos with the United Kingdom and United States to dismantle criminal infrastructure	Active + reactive
2024	DFAT, **ACSC**	Leading role in issuing an international security advisory and guidelines on APT 40	Reactive

Source: Author's analysis.

SINGAPORE

Strategy and Governance

In the **Singapore Cybersecurity Strategy 2021**, Singapore announced plans to adopt an active approach to risk identification, risk management, and the defense of its networks. However, Singapore's "proactive response" differs from the approaches of the United States and the United Kingdom to ACD; it focuses not on inflicting direct effects on adversaries but on prevention, early detection, information sharing, and regulatory strengthening. Accordingly, Singapore has incorporated a proactive, preventive approach into its cybersecurity **regulations** and has imposed stricter legal obligations and incident-reporting requirements on critical infrastructure operators.

The Cyber Security Agency (CSA), under the prime minister's office, serves as Singapore's national cybersecurity authority. In October 2025, the CSA established the **Digital Defence Hub** (DDH) under the Centre for Strategic Infocomm Technologies (CSIT), assigning it missions such as threat hunting, malware analysis, and red team exercises to detect and prevent cyber threats before they cause harm. However, these activities remain focused on preventive and defensive measures.

In addition, the Digital and Intelligence Service (DIS) within the Ministry of Defence of Singapore (MINDEF) leads cyber operations, protects the networks of the Singapore Armed Forces (SAF), and contributes to national cyber resilience. For this purpose, MINDEF established the **Defence Cyber Organisation** (DCO). In addition, Singapore announced on March 3, 2025, plans to restructure DIS into a **Defense Cyber Command** (DCCOM) and Digitalisation Command, with the DCCOM expected to enhance cybersecurity and operational capabilities. However, the SAF's cyber capabilities remain primarily oriented toward defensive functions such as network intrusion detection, protection, and recovery. A 2023 **report** by the International Institute for Strategic

Studies noted that although Singapore acknowledges the need for offensive cyber capabilities, concrete discussions about their development remain limited.

Singapore, however, places greater emphasis on reactive measures. It hosts INTERPOL's **Innovation Centre** and actively participates in joint international responses to global cybercrime. The **Singapore Police Force** (SPF) leads cybercrime investigations and cooperation with INTERPOL. As a result, Singapore primarily engages in indirect threat removal through participation in international cooperative operations rather than through independent extraterritorial cyber actions. This approach also appears in **Singapore's attribution practices**. Singapore prefers technical attribution over politically charged public accusations against states.

Singapore also works to strengthen cybersecurity capacity across the Association of Southeast Asian Nations (ASEAN) to close regional security gaps and enhance the overall trustworthiness of regional networks. It leads initiatives such as the ASEAN Cyber Capacity Programme (ACCP) and the **Cyber Capabilities and Capacity Project** (C3DP) within INTERPOL, helping build cyber and law enforcement capabilities in neighboring countries. Singapore's foreign minister has **emphasized** that "cybersecurity is only as strong as its weakest link," underscoring the importance of global cooperation to enhance cybersecurity capabilities, particularly in developing countries.

Singapore Cases

Recent cases illustrate Singapore's active participation in law enforcement cooperation with INTERPOL, as well as the country's increasing engagement in joint responses to cyber incidents. In July 2025, Singapore publicly **attributed**—for the first time—a cyber threat actor linked to China. The minister for home affairs accused the actor known as UNC3886, following Mandiant's naming convention, of conducting an advanced persistent threat campaign targeting critical infrastructure in Singapore. However, the country did not directly name China and avoided using labels tied to Chinese government entities such as Volt Typhoon, instead relying on an unnamed cluster group designation to deliver an indirect message. This approach reflects the position of a small country with deep economic, regional, and geopolitical ties to China—a strategy of condemning malicious activity and asserting sovereignty while minimizing diplomatic conflict or risk of retaliation.

Singapore has also cooperated with **Thailand**, Hong Kong, **Malaysia**, and others in joint operations targeting Southeast Asian cyber fraud networks, resulting in the arrest of perpetrators, takedown of infrastructure, and seizure and freezing of illicit assets.

Table 5: Singapore Cyber Response Cases

Year	Entities	Response	Type
2023	**SPF**	Conducting Operation Synergia with INTERPOL; takedown of servers; arrests of suspects	Active + reactive
2023	**SPF**	Working with INTERPOL to take down a phishing-as-a-service (PhaaS) platform; arrests of suspects	Active + reactive

2024	**CSA**	Participation in internationally coordinated Operation Secure; disruption and disabling of a global botnet infrastructure; neutralization of approximately 2,700 infected devices in Singapore	Active
2025	**MHA**	Public attribution and condemnation of a cyberattack on national critical infrastructure to state-sponsored APT group UNC3886	Reactive
2025	**SPF**	Investigation of the Cambodia-based online fraud group Prince Group; freezing and seizure of approximately SGD 150 million in assets located in Singapore	Reactive

Source: Author's analysis.

SUMMARY AND IMPLICATIONS FOR SOUTH KOREA

Major countries' cyber defense approaches increasingly extend beyond traditional passive defense to include more active measures for detecting, disrupting, and blocking threats. While national approaches vary, they generally share a common structure that combines elements of ACD with **reactive** mechanisms.

The United States represents one of the more forward-leaning ACD models. Through **the defend forward** and **hunt forward concepts**, the United States conducts ACD operations beyond its own networks to detect and disrupt threats from foreign environments. This approach seeks to reduce an adversary's operational space and shape behavior through **persistent engagement** in cyberspace. After an attack, agencies such as the FBI, USDT, and DOS coordinate criminal indictments, sanctions, and diplomatic attribution, forming an integrated system of legal, economic, and diplomatic cost imposition. The United States also extends its cyber defense posture into a cooperative framework with allies and partners, emphasizing information sharing and collective defense and deterrence.

The United Kingdom conducts offensive cyber operations through the NCF, a combined entity of the intelligence community and the military, and officially acknowledges using these capabilities on a "daily and routine" basis. The country clearly classifies such actions as **offensive cyber operations**, while emphasizing legitimacy and transparency by adhering to principles of proportionality, necessity, and accountability under domestic and international law. The United Kingdom actively uses sanctions—supported by its independent cyber sanctions framework—and conducts extensive international coordination and joint responses alongside the United States.

Japan recently codified the concept of ACD into law and has actively reorganized its institutions to implement this shift. The NCO serves as the central authority for directing and coordinating ACD, while operational activities are carried out by the JSDF and NPA. An independent oversight body conducts prior review and approval of ACD measures and ensures the protection of communications privacy and personal data. This system is designed to ensure that controversial

measures are conducted through legal procedures, thereby securing both legality and public acceptability. Japan has also begun to take a more active posture in public attribution, including direct attribution of cyberattacks to China.

Australia designates offensive cyber operations as an explicit mandate of the **ASD** and maintains distinct authorization and oversight procedures depending on operational purpose. The ASD operates a standing joint framework with the AFP and supports military cyber operations. Australia is currently making large-scale investments to strengthen the ASD's cyber capabilities and has expanded its reactive posture by pursuing joint sanctions and attributions with partners such as the United States and United Kingdom.

Singapore, as a relatively small state, adopts resilience building and active international cooperation as core strategies to avoid unnecessary geopolitical conflict. It strengthens proactive measures through early threat detection, information sharing, and recovery capabilities, while calibrating public attribution cautiously. At the same time, Singapore clearly signals its national will to respond to cyberattacks and seeks to enhance joint cyber response efforts through multilateral cooperation (e.g., via ASEAN and INTERPOL) as a means of securing deterrent effects.

Across these cases, the designation of "offensive" and "defensive" cyber operations is not uniform. Some states, such as the United Kingdom and Australia, define cyber operations as offensive acts, while others, such as the United States and Japan, frame them as extensions of defensive measures. Despite these differences in terminology and framing, all countries emphasize compliance with domestic and international legal norms, particularly proportionality, necessity, and accountability. To support these approaches, governments are increasing budgetary allocations and investing in the development of specialized personnel and operational capabilities.

Accordingly, South Korea should recalibrate its cyber defense strategy within an integrated framework that combines ACD with reactive response mechanisms. This should not be understood as offensive action aimed at attacking an adversary, but rather as a defensive approach that includes actively identifying and disrupting threats, as well as undertaking reactive measures and resilience-building efforts after an incident. What is needed, therefore, is a coherent strategy in which ACD and reactive response elements operate in a complementary manner, supported by sustained efforts to build the capabilities required to implement them effectively.

Current Status and Directions for Improvement in South Korea

SOUTH KOREA'S CYBER DEFENSE

Strategy and Governance

South Korea first introduced the term "active response" in its 2019 **National Cybersecurity Strategy**. Although the strategy did not define active defense, it included, under the task of developing "comprehensive and active response measures," efforts to prepare response tools consistent with international norms and expand cyber warfare personnel in the event of a significant cyber threat.

The 2024 **strategy** replaced the term "active" with "proactive" and "offensive," signaling the government's intention to move away from a predominantly defensive and reactive posture. The strategy identified "strengthening offensive cyber defense activities" as a major task and outlined detailed measures such as expanding attribution and responsibility assignment, establishing a deterrence strategy, blocking threats proactively, and developing diplomatic and technical tools to counter influence operations.

South Korea has also sought to establish a unified national cyber threat response system centered on the National Intelligence Service (NIS) and the National Security Office (NSO). In 2023, the government established the **National Cyber Crisis Management** Center, led by the NIS and operating under the oversight of the NSO, with participation from relevant ministries.

The legal and institutional basis for South Korea's ACD was largely established through the **National Intelligence Service Act**, in a way similar to the United Kingdom and Australia. The act designates "prevention and response to cyberattacks and threats" as an NIS mandate. This reflects the North Korea–centric threat environment facing South Korea and the fact that North Korea's early cyber activities were more aligned with espionage and asymmetric attacks than with military operations, and were thus addressed by the traditional counterintelligence process at the NIS. The **Regulation on Cybersecurity Service**s, a subordinate regulation under the NIS Act, explicitly authorizes the NIS to conduct active cyber defense measures, including the tracking and neutralization of infrastructure located abroad.

As a military organization, the Cyber Command was established in 2011 and reorganized in 2019 into the **Cyber Operations Command** under the Joint Chiefs of Staff (JCS), following the removal of its psychological operations function. North Korea–related **psychological operations** are now performed by the Psychological Operations Group under the Ministry of National Defense (MND). The **Defense Counterintelligence Command** (DCC) supports the military's cyber protection posture and contributes to information operations.

The Cyber Operations Command is responsible for a wide range of tasks, including cyber operations planning and execution, security activities, system development and maintenance, training, information sharing, and threat intelligence collection. Since 2019, each military service—the **army**, **navy**, and **air force**—has established its own Cyber Operations Center responsible for defensive operations and security monitoring. In 2023, the NIS's National Cybersecurity Center (NCSC) and the Cyber Operations Command signed an **MOU** to formalize interagency cooperation.

Public attribution in South Korea is carried out by multiple agencies, including the NIS, National Police Agency (NPA), and Ministry of Foreign Affairs (MOFA), while cybersecurity advisories are primarily issued by the NIS. Investigations and law enforcement related to cyber-enabled national security crimes are led by the NPA and the Supreme Prosecutors' Office (SPO). Although relatively limited in number, sanctions—particularly those related to North Korea—are imposed by the MOFA and enforced by the Ministry of Economy and Finance (MOEF) and the Financial Services Commission (FSC).

South Korea Cases

South Korea's responses to cyberattacks have been focused predominantly on North Korea. Until recently, responses to North Korean cyberattacks have consisted mainly of public attribution and statements of condemnation. However, in 2023, South Korea imposed its first independent **sanctions** related to North Korean cyber activities, signaling efforts to shift toward a more active response posture. Recently, attempts have been made to expand responses into coordinated, multiagency actions involving the NIS, NPA, MOFA, and MOEF.

Table 6: South Korea Cyber Response Cases

Year	Entities	Response	Type
2019	**NPA**	Public attribution and condemnation of the Upbit cryptocurrency theft	Reactive
2023	NIS, **NPA**, **MOFA**, MOEF	Joint attribution and sanctions against the Kimsuky hacking group	Reactive
2023	**NPA**	Public attribution and condemnation of cyber activities targeting ROK–U.S. joint military exercises	Reactive
2023	**NIS,** NPA, **MOFA**	Joint attribution and condemnation of North Korean IT workers; sanctions	Reactive
2023	MOFA, **MOEF**	Joint attribution and condemnation of illicit cyber activities involving technology theft; sanctions	Reactive
2024	**NPA**	Public attribution and condemnation of a cyberattack on the court network	Reactive
2024	**NPA**, NIS	Joint attribution and condemnation of cyber activities targeting weapons and military-related technologies; security advisory; international cooperation	Reactive
2024	**NIS**, NPA, SPO, DCC, Cyber Operations Command	Public attribution and condemnation of the attempted theft of defense technology; security advisory	Reactive
2024	MOFA, **MOEF**	Joint attribution and condemnation of North Korean IT workers; sanctions; international cooperation	Reactive
2025	**NPA**	Participation in Operation HAECHI targeting cyber financial crime; sanctions; freezing and seizure of assets; cooperation with INTERPOL	Reactive

Source: Author's analysis.

CHALLENGES AND CONSTRAINTS

Lack of Substantive Post-Attribution Measures

South Korea has experienced persistent cyberattacks from North Korea and has repeatedly issued attribution statements and condemnations. However, legal, economic, diplomatic, and ACD follow-up measures have lacked consistency and specificity.

In the international community, naming and shaming has been widely used as a response option to cyber threats, with an increasing number of states participating in this practice. Yet in many cases, the attributed state simply denies responsibility, and attribution is not followed by substantive measures, thus reducing its effectiveness and credibility. In South Korea, repeated attribution without meaningful consequences risks being perceived as either limited capability or insufficient political will.

This tendency is closely linked to broader strategic considerations. In managing relations with North Korea, the South Korean government has often prioritized nuclear security, crisis management, and inter-Korean stability over assertive cyber responses. Even when cyber operations do not directly escalate military tensions, concerns about destabilizing the broader security environment may lead to calibrated and restrained responses. As a result, cyber policy has at times reflected strategic caution rather than sustained cost imposition.

The 2024 National Cybersecurity Strategy signaled a shift toward a more forward-leaning posture by adopting the language of "offensive" cyber defense. However, publicly observable implementation has remained limited. Although economic sanctions and regulatory measures targeting cryptocurrency exchanges were introduced to curb North Korea's illicit financial gains, their impact has been constrained. North Korean actors have increasingly adapted by shifting to more opaque financial networks, including informal brokers and cryptocurrency mixers operating through **China** and **Russia**, thereby reducing the effectiveness of existing measures.

Insufficient Oversight and Authorization Mechanisms for ACD

Major countries that conduct offensive or ACD operations have established clear domestic legal bases as well as pre-authorization and oversight procedures, emphasizing legitimacy and transparency. In South Korea, although the NIS is granted authority for ACD, the legal basis remains a broad delegation framed as "may take necessary measures, such as tracking and neutralizing."

Under current cybersecurity regulations, this authority allows the director of the NIS to identify, deter, and block foreign-based hacking infrastructure that threatens national security or national interest. However, the procedures for authorization, the structure of oversight, and the allocation of responsibility for failures or misuse are not sufficiently specified. The NSO is designated as the control tower for national cybersecurity, but its concrete role in reviewing and approving NIS's ACD measures is neither clearly defined nor operationalized.

This creates dual constraints: On one hand, the NIS director's discretion may be interpreted too broadly; on the other hand, legal and political liability concerns may discourage the agency from employing ACD capabilities. Without an oversight and authorization framework that ensures

democratic control and transparency, the legitimacy and trustworthiness of ACD operations remain vulnerable to challenge, hindering their practical use.

Limitations on Military Roles and Authorities

Constraints also exist within the military domain. Although placing the Cyber Operations Command under the JCS reflects an intention to link cyberspace with broader military operations, the traditional mission set of the JCS—focused on land, sea, and air forces—raises the risk that cyber operations may be treated as auxiliary rather than as an independent strategic domain.

Moreover, because cyber operations blur the boundaries between peacetime and wartime, requiring persistent engagement and reconnaissance in the gray zone, the extent to which Cyber Operations Command can conduct ACD measures during peacetime is not clearly established.

Repeated political controversies—such as the **2013** online comment manipulation scandal and Cyber Operation Command's alleged involvement in intelligence activities during the **2024** martial law-related mobilization incident—have also weakened trust in the organization. These controversies led to the **removal** of its psychological operations function in 2019 and reinforced a risk-averse internal culture. This environment discourages the military from more actively implementing ACD. South Korea's constrained military posture therefore stands apart from emerging trends in other countries that are expanding dedicated personnel and capabilities for military cyber operations, and this limitation restricts the military's ability to assume a more active role in ACD.

Insufficient Interagency Cooperation Mechanisms

Gaps remain between the formal interagency cooperation framework and its actual implementation. Although agencies such as the NIS, MND, NPA, MOFA, and financial authorities each perform roles in cyber incident response, it is rare for technical analysis, investigation, prosecution, sanctions, diplomatic action, and ACD measures to form a coordinated set of response measures for a given cyber campaign.

There are also inconsistencies as to which agencies participate in a given case (reflected in Table 6). For example, despite joint investigative coordination by the National Cyber Crisis Management Center regarding North Korea's cyber activities during the 2023 ROK-U.S. joint military exercise, the final announcement was issued solely by the NPA. This may be because the NIS primarily issues security advisories, while the MOFA tends to accompany such cases with sanctions or other diplomatic measures; this resulted in the NPA taking the lead in announcing the investigation results. However, such single-agency announcements make it difficult to view the government's actions as a genuinely integrated national response to North Korea's cyber threats.

In the 2024 defense industry hacking case, the NIS issued a **security advisory**, but it was only months later that the police and the Defense Acquisition Program Administration jointly announced follow-up inspections, which were once again limited to **public attribution** without further substantive measures. Compared with the U.S. model—in which the NSA, CISA, FBI, DOJ, USDT, and DOS coordinate multilayered responses—South Korea's fragmented structure is a limitation.

The 2023 MOU between the NIS and Cyber Operations Command is a step toward closer cooperation, but its necessity highlights the absence of an institutionalized information-sharing and operational coordination system. The two institutions operate under different legal authorities and regulatory frameworks, and they also differ in security classification practices and operational approaches. As a result, institutionalized information sharing and personnel exchanges are difficult to implement. Given the declaratory nature and limited legal force of the MOU, structural constraints on sustained cooperation remain.

Divergent Threat Perceptions and the Limitations of North Korea's Shift to Overseas Operations

North Korea's cyber operations increasingly target foreign financial institutions, cryptocurrency exchanges, global IT companies, and third-country infrastructure rather than South Korea directly. According to the **2025 Microsoft Digital Defense Report**, only 1 percent of observed North Korean cyber operations targeted South Korea—an especially sharp decline compared to **2023**, when attacks against the United States and South Korea accounted for 50 percent of North Korea's cyber operations. This indicates a structural shift toward overseas operations aimed at acquiring funds, technology, and infrastructure. It creates fundamental constraints for South Korea, as direct control or legal jurisdiction over third-country environments is limited.

However, reduced direct targeting does not imply reduced relevance. Funds and technology acquired overseas are funneled into North Korea's nuclear and missile programs, ultimately affecting South Korea's security environment. Attacks occur abroad, but their consequences cycle back to South Korea.

Moreover, South Korea's international cooperation remains narrowly focused on North Korea, even as China's and Russia's cyber activities increasingly impact South Korea's economic and technological security. Geopolitical and economic sensitivities further constrain South Korea's ability to take direct or unilateral action against major powers.

Policy Recommendations

EXPANDING PRACTICAL RESPONSES AND BUILDING AN OPERATIONAL TRACK RECORD

South Korea should expand substantive post-attribution responses within an integrated framework. Attribution should be one of the key elements feeding into a broader integrated response—one that brings together four elements: (1) legal measures (e.g., criminal indictments and joint investigations); (2) economic sanctions (e.g., asset freezes and financial restrictions); (3) diplomatic action (e.g., joint attribution, condemnation, and multilateral forums); and (4) technical measures (e.g., disrupting or dismantling malicious infrastructures, including those located overseas).

Although such measures may not directly alter North Korea's behavior—given its existing isolation and limited sensitivity to reputational costs—they can increase the operational and attack costs on third-country infrastructure, facilitators, financial channels, and IT personnel supporting North Korean operations. Repeated constraints on third-country enablers can raise the risk and burden of maintaining these operations, ultimately reducing their efficiency and impact.

The South Korean government should develop its own response options and strengthen its capabilities to implement both ACD and reactive response measures. Through sustained and tangible implementation, South Korea can build a track record of integrated response that reinforces credibility.

CLARIFYING LEGAL PROCEDURES AND OVERSIGHT FOR ACD

To conduct ACD operations, South Korea should first establish clear domestic authorization and oversight frameworks. This requires defining the roles of the president and NSO, NIS and MND, and the National Assembly. A multilayered structure should include: (1) top-level authorization, (2) delegated execution with internal controls, and (3) external democratic oversight.

Significant actions with military or diplomatic implications—such as operations affecting foreign critical infrastructure—should require prior approval from the president or NSO. The NSO could ensure that a substantive review is carried out, including an assessment of necessity, proportionality, and domestic and international legality.

Limited ACD operations may remain under the authority of the NIS director, but with clearer scope and conditions. The NIS could execute such measures swiftly while reporting outcomes to the NSO. Distinguishing between presidential/NSO authorization and delegated authority would help balance responsiveness with control.

Oversight should combine internal and external mechanisms. Internally, the NSO could maintain records and conduct reviews of all ACD activities. Externally, regular classified reporting to relevant National Assembly committees could provide minimum democratic accountability.

DEFINING ROLES ACROSS PEACETIME, CRISIS, AND WARTIME AND STRENGTHENING MILITARY CYBER CAPABILITIES

An integrated cyber response framework should establish sequential role division across peacetime, crisis, and wartime. Intelligence agencies possess comparative advantages in tracking North Korean espionage-oriented cyber threats and overseas infrastructure during peacetime and periods of gray zone confrontation, while the military should lead combined and joint cyber operations during wartime.

In peacetime, the NIS should lead persistent engagement while the military provides operational support. During crisis periods, the National Cyber Crisis Management Center should direct, coordinate, and control the full spectrum of cyber defense activities, with the NIS and MND cooperating under its direction. In wartime, operational control for cyber offense and defense should transition to the military, with intelligence agencies providing continued support. Relevant authorities, missions, and command structures could be codified in law or operational manuals, with clear transition procedures across phases.

Military active and offensive cyber capabilities should also be strengthened. This includes insulating the Cyber Operations Command from political controversies, reinforcing operational units dedicated to attack, active defense, and joint operations, and enabling focus on disruptive actions against adversary infrastructure and military cyber effects. Moreover, actual operational execution

requires cultivating legal specialists capable of assessing risks and advising on the lawful conduct of cyber operations. In the long term, limited credible signaling of offensive capabilities may be considered to enhance deterrence.

INTEGRATED USE OF CYBER, LEGAL, DIPLOMATIC, AND ECONOMIC TOOLS (WHOLE-OF-GOVERNMENT RESPONSE)

Effective cyber response requires integrated action rather than isolated, agency-specific measures. Technical analysis, investigative action, prosecution, sanctions, diplomatic actions, and ACD measures should be linked as a unified response to specific attacks or campaigns.

A standardized joint response process should link actions by the NIS, MND, NPA, MOFA, MOEF, and others. The National Cyber Crisis Management Center should strengthen its coordinating function so that technical assessments, disruption efforts, investigative direction, diplomatic action, and sanctions decisions are not merely shared but strategically integrated. An integrated architecture would allow South Korea to convey consistent messages and impose meaningful costs on malicious actors.

EXPANDING INTERNATIONAL COOPERATION BEYOND A NORTH KOREA–CENTRIC APPROACH

Given that North Korea conducts most of its cyber operations abroad, the room for unilateral South Korean response is inherently limited. In light of this reality, South Korea should gradually expand its extraterritorial ACD activities. To ensure that such actions remain within the bounds of international law, however, the South Korean government should carry them out with the prior consent and cooperation of relevant third-country authorities or through joint operations with the United States and other allies. Building these conditions requires a broad framework of international cooperation that includes joint action, investigative cooperation, capacity-building support, and information sharing with allies and partners.

Moreover, considering the growing impact of Chinese and Russian cyber activities on South Korea's economic and technological security, South Korea should expand its threat perceptions beyond North Korea and increase participation in multilateral and international responses. Such engagement not only builds trust with allied partners and provides an indirect means of responding to Chinese and Russian cyber activities, but also helps strengthen the diplomatic and operational foundations for more coordinated international responses to North Korea's cyber threats.

Cooperation with third countries whose infrastructure and financial networks are exploited by North Korean actors, and where North Korean personnel operate, is also crucial. Recent initiatives—such as the **MOU** with China on voice phishing and online fraud, or the establishment of an **international cooperation mechanism** for cross-border scam response—represent meaningful progress. Providing capacity-building assistance to strengthen the cyber defenses of third countries not only disrupts North Korea's revenue-generation channels but also helps secure third countries' consent and cooperation for expanding ACD measures against North Korean operations conducted on their territory. In particular, given China's role as a key transit environment for North Korean

cyber activities—including the movement, laundering, and cash-out of illicitly obtained funds—cooperation with China remains indispensable for effectively addressing these threats.

Ultimately, South Korea should evolve into an active and meaningful contributor to international cyber response efforts. By expanding international cooperation along these lines, South Korea can enhance its ability to counter overseas cyber threats.

Sunha Bae *is a senior researcher on the Cybersecurity Policy Research Team at the National Security Research Institute of Korea.*

Cross-Border Law Enforcement Collaboration for Countering North Korea's Crypto Plunder

By Joohui Park

Introduction

The Democratic People's Republic of Korea (DPRK) is exploiting cyberspace and decentralized cryptocurrency architecture in order to satisfy its craving for crypto assets, and it has become a global security concern.[1] This is reflected in the **final report** adopted in 2025 by the UN Open-Ended Working Group on "security of and in the use of information and communications technologies 2021-2025," which acknowledges growing member-state concerns that cryptocurrency theft may affect international peace and security. These concerns, expressed throughout the working group process, were incorporated into the collective threat assessment contained in the final report. Later in 2025, the Multilateral Sanctions Monitoring Team (MSMT) issued **recommendations** in response to violations of UN Security Council resolutions by North Korea, particularly regarding cryptocurrency theft and laundering practices used to circumvent sanctions.

North Korea's cryptocurrency theft and subsequent money-laundering activities are highly distinctive. First, they are conducted by state-directed criminal networks, often enabled by overseas-deployed information technology (IT) workers. Second, these activities rely on cryptocurrency laundering tools from third-party service providers outside of North Korea. Finally,

facilitators operating in third countries covertly enable the conversion of stolen cryptocurrency into fiat currency.

To counter this emerging threat posed by North Korea, South Korea has undertaken a comprehensive set of countermeasures. These include strengthening **domestic regulation** of virtual-asset service providers, notably through (1) the introduction of **know-your-customer (KYC) and information sharing obligations** for cryptocurrency exchanges, (2) imposing targeted **sanctions** on individuals and entities involved in cryptocurrency theft and money-laundering activities, (3) enhancing **law enforcement cooperation** with domestic and foreign cryptocurrency exchanges to disrupt laundering flows, and (4) expanding international cooperation, particularly with **the United States and Japan** and **the Financial Action Task Force (FATF)**, an organization discussed in more detail later in this paper.

When it comes to enforcing laws against North Korea's crypto-related illicit activities, however, South Korea continues to face significant challenges. First, the inherently decentralized nature of virtual asset transactions complicates efforts by law enforcement authorities to identify and trace perpetrators involved in cryptocurrency theft. Second, uneven or absent international regulatory frameworks governing virtual-asset finance impede the ability of law enforcement agencies to obtain timely cooperation from foreign cryptocurrency exchanges. Third, the lack of an internationally coordinated framework for law enforcement cooperation in the investigation, freezing, seizure, or confiscation of proceeds derived from cryptocurrency theft constrains the ability of national authorities to secure swift cross-border assistance. Against this backdrop, this paper analyzes the potential of the UN Convention against Cybercrime (UNCC) as an international law enforcement framework for countering North Korea's crypto theft and laundering practices.

North Korea's Craving for Crypto Assets

North Korea's pursuit of cryptocurrency reflects a broader strategic adaptation to international sanctions and financial isolation. As conventional revenue streams were severed, the regime turned to cyberspace—exploiting decentralized crypto infrastructure to generate funds beyond the reach of traditional regulatory frameworks. This section examines why the regime shifted its cyber operations toward large-scale crypto theft, and the methods it employs to launder and cash out stolen assets.

A SHIFT IN PRIORITIES: THEFT TO SUSTAIN THE REGIME

Over the past decade, North Korea has systematically exploited cyberspace to acquire assets—both cryptocurrency and fiat currency—to finance its weapons programs and nuclear ambitions. Prior to this period, North Korea engaged in malicious cyber activities for objectives other than asset acquisition. A notable example is the **2013 Dark Seoul attack**, which targeted South Korean financial institutions and media outlets, causing extensive damage to data and information systems. The primary objective of the attack was to instill public panic through the disruption of critical services, although it also resulted in economic losses—approximately **48,700 PCs and servers** were disrupted, with estimated **losses of KRW 867.2 billion**.

Since the 2016 adoption of **UN Security Council Resolution 2270**, which imposed comprehensive sanctions on North Korea following its fourth nuclear test, the regime's primary use of cyberspace has shifted markedly toward the theft of funds. For example, in early 2016, DPRK actors **attempted** to steal fiat currency by exploiting vulnerabilities in the Bangladeshi central bank's SWIFT system. At that time, the use of virtual assets was still nascent and had not yet become a viable vehicle for large-scale illicit finance. However, following the **complete exclusion** of DPRK banks from the SWIFT network in March 2017, the country lost all formal channels for transmitting or receiving U.S. dollars. Against this backdrop, cryptocurrency theft has emerged as a critical mechanism through which North Korea generates revenue while circumventing conventional financial regulations. While North Korea appears to continue maintaining units responsible for disruptive and destructive cyber operations, even these units are reportedly required to be **self-financing**, which in practice drives them to engage in cybercrime as a means of generating revenue for the regime.

The scale of cryptocurrency theft attributed to DPRK actors has become increasingly alarming. According to **a 2025 report** by the MSMT, North Korea stole approximately USD 1.19 billion in cryptocurrency in 2024 and at least USD 1.645 billion between January and September 2025. Notably, these figures do not include the alleged theft of an additional USD 30 million from **Upbit**, a major cryptocurrency exchange based in South Korea, in late 2025. Moreover, the value of crypto assets stolen in single incidents continues to reach unprecedented levels. In February 2025, North Korean cyber actors allegedly stole nearly **USD 1.5 billion** in Ethereum tokens (ETH) from Bybit, the world's second-largest cryptocurrency exchange. This incident constitutes **the largest cryptocurrency theft** on record, surpassing the 2022 Ronin Bridge hack, which had previously held that distinction with losses estimated at approximately **USD 620 million**.

MODUS OPERANDI: THE 2022 RONIN NETWORK THEFT

North Korea's crypto theft and subsequent money-laundering process are well-coordinated by its IT workers, resourced by crypto-laundering tools, and assisted by facilitators within safe havens. North Korea mobilizes IT workers abroad, masquerading as legitimate employers or employees, to implant malicious code. Diverse third-party crypto-laundering tools, such as mixing and bridging services, are used to obfuscate tracing. Further, facilitators in third countries are contributing to the conversion of stolen cryptocurrencies into fiat currencies. In the course of an attack, North Korean actors commit numerous types of criminal activities, including identity theft, cybercrime (e.g., illegal access and misuse of tools), and money laundering.

To understand the challenges in law enforcement in relation to these criminal activities by North Korean actors, it is worth following North Korea's primary patterns of behavior in crypto thefts and money laundering. The Ronin Network theft in 2022 represents a general playbook carried out by North Korea for crypto theft and laundering.

The first step is cyber-enabled theft. North Korean cyber operators deploy the full spectrum of available methods to obtain assets, including cutting-edge technical capabilities (e.g., AI) and refined social engineering tactics. In early 2022, employees at Sky Mavis, the developer of Axie Infinity game, were targeted by threat actors impersonating recruiters of a fictitious company who solicited job applications through LinkedIn. A fraudulent job offer was then delivered via a

weaponized PDF document. Upon download, the file deployed malware that infiltrated the Ronin Network's bridge infrastructure, which enables cross-chain asset transfers for Axie Infinity. This initial compromise enabled the attackers to gain control of **four of the nine validators** securing the Ronin Network, leaving them one validator short of achieving complete network control. Using these validators, the actors authorized two fraudulent withdrawal transactions: **173,600 ETH and 25.5 million USD Coin**, equivalent to more than USD 600 million.

The second step is cryptocurrency laundering using a multistage methodology. Once stolen cryptocurrency is located in wallets under their control, DPRK actors initiate an extensive money laundering operation, using a diverse array of tools such as mixing, swapping, and bridging.[2] According to Chainalysis, the laundering scheme has utilized over **12,000 distinct crypto addresses**, demonstrating the group's highly sophisticated obfuscation capabilities. In the Ronin case, the cyber operators first sent stolen crypto assets to intermediary wallets. Subsequently, crypto assets were mixed in batches through Tornado Cash, a decentralized mixing service which contributed to laundering over **USD 455 million** worth of crypto stolen from this Ronin Network theft. Obfuscated crypto assets in ETH were swapped for bitcoin (BTC) to exploit inter-blockchain tracing complexities. The crypto assets in BTC underwent additional batch mixing to further obscure transaction provenance.

The final stage is the "cash out" phase, which constitutes the critical end goal for DPRK operators converting illicit cryptocurrency holdings into liquid fiat currency that can fund state activities. In most cases, hackers enter a dormant state, waiting for suitable timing to cash out. For six months, a large portion of the crypto assets stolen from the Ronin Network heist remained unmoved in cryptocurrency wallets under the DPRK actors' control. With the collaboration of law enforcement authorities and Chainalysis, **10 percent** of stolen funds were recovered in September 2022.

The Hide-and-Seek Chase for Crypto Theft

From a law enforcement perspective, investigating, prosecuting, and recovering assets from cryptocurrency theft attributed to North Korean actors presents significant challenges. This section examines the key obstacles confronting law enforcement, from the decentralized and pseudonymous architecture of the cryptocurrency ecosystem to the exploitation of third-country facilitators, as well as the steps South Korea has taken to confront these challenges.

THE DECENTRALIZED AND PSEUDONYMOUS ARCHITECTURE

The principal challenges in investigating and prosecuting criminal activities associated with cryptocurrency theft stem from the structural features of the cryptocurrency ecosystem itself. Cybercrime, in most cases, is inherently transnational in nature and therefore already presents significant obstacles for cross-border law enforcement. Beyond this general difficulty, however, cybercrimes involving cryptocurrencies give rise to distinct challenges derived from two defining characteristics of the crypto ecosystem: the decentralized and pseudonymous nature of transactions.

First, cryptocurrency transactions are decentralized by design. Cryptocurrencies function as a medium of exchange whose issuance, transfer, and ownership are recorded through distributed ledger technologies, most notably blockchain. For law enforcement authorities, such distributed ledgers constitute a double-edged sword. On the one hand, every cryptocurrency transaction is immutably recorded on a publicly accessible ledger. This transparency enables investigative authorities to trace and analyze financial flows with a degree of precision. On the other hand, transactions recorded on distributed ledgers do not intrinsically require intermediaries. This does not mean that there exist no intermediaries; of course, intermediaries such as crypto exchanges exist for other purposes, including facilitating cryptocurrency transactions. However, transactions based on blockchain do not by design necessitate any intermediaries as a prerequisite for concluding transactions. With only wallet addresses, parties can transact directly on a peer-to-peer basis. In such transactions, law enforcement can hardly expect the cooperation they have with regard to centralized transactions—for instance, identity verification of transaction parties—which are conducted under the control of traditional banks. Moreover, stolen crypto assets may be stored in unhosted wallets. Because only the wallet holder can exercise control over assets stored in such wallets, law enforcement authorities are **unable to freeze, seize, or recover** these funds in the absence of access to the relevant private keys.

Second, cryptocurrency transactions are **pseudonymous**. Wallet addresses required to conduct transactions consist merely of alphanumeric strings, and the verification of a user's real-world identity is not a prerequisite for participation in the network. As a result, while investigative authorities may be able to trace the movement of cryptocurrency across the blockchain, they are unable to ascertain the identity of the individuals or entities controlling the relevant wallets, or it is extremely challenging to do so.

These architectural features of cryptocurrencies significantly enhance North Korea's incentives to target crypto assets, as they allow state-sponsored actors to evade detection or, at minimum, delay and complicate tracing efforts in the cyber domain. Such incentives are further amplified by the development of sophisticated obfuscation techniques—including mixing, swapping, and cross-chain bridging—designed to limit or effectively eliminate transactional traceability. In addition, the emergence of **anonymity-enhanced cryptocurrencies**, which are specifically engineered to obscure transactional linkages between wallet addresses and thereby undermine blockchain analytics, further impedes law enforcement efforts.

Consequently, blockchain-specific technical expertise has become indispensable for effective law enforcement. Cross-border cooperation in crypto-related investigations frequently depends on a requesting authority's ability to present credible, technically substantiated evidence demonstrating that stolen assets are located within the jurisdiction of the requested state. In practice, however, most law enforcement agencies lack the specialized expertise required to conduct such analyses independently. While assistance from private blockchain analytics firms—such as Chainalysis and Elliptic—remains valuable, investigative authorities must also develop and institutionalize in-house cryptocurrency and blockchain analytical capabilities to ensure sustainable and effective enforcement.

TARGETING CASH-OUT POINTS

Although the cryptocurrency ecosystem is predominantly decentralized, centralized exchanges (CEXs) exist to facilitate transactional convenience for users.[3] They serve as fiat-to-crypto gateways that enable users to trade fiat currencies for virtual assets with ease, and vice versa. Also, they provide user-friendly asset management for those unable or unwilling to manage private keys independently and play a critical role in providing a wide range of crypto-related services.

Criminal actors exploit these CEXs primarily to convert crypto assets into fiat currency. Consequently, law enforcement agencies have targeted these cash-out points as nodes for tracing, freezing, seizing, and confiscating stolen cryptocurrency. In several jurisdictions, cryptocurrency CEXs are subject to **anti-money laundering and counter-financing of terrorism** obligations, including KYC requirements comparable to those imposed on traditional banks and financial institutions.

By capturing the moment when DPRK actors use crypto exchanges to cash out stolen money, it is possible to identify, freeze, seize, or even recover the stolen crypto money. For example, the approximately 10 percent of stolen funds recovered from the 2022 Ronin Bridge heist were seized by targeting cash-out points within the cryptocurrency ecosystem. This operation marked the first recorded seizure of stolen cryptocurrency, demonstrating the feasibility of asset recovery through law enforcement intervention at centralized exchange chokepoints. Blockchain analysis provided by Chainalysis was used to trace the stolen funds to **specific cash-out points**.

In another example, a portion of the cryptocurrency stolen in **the 2019 Upbit heist** was recovered in 2024. In that incident, the exchange suffered the theft of approximately USD 41 million ETH. This case represents the first substantiated instance in which a South Korean investigative authority formally attributed the theft of cryptocurrency from a domestic exchange to DPRK operatives. The South Korean Police Agency (SKPA) worked closely with the FBI and **concluded** that the heist had been perpetrated by DPRK actors. The SKPA ultimately recovered approximately 4.8 BTC, valued at roughly USD 400,000, which were subsequently returned to Upbit.

However, the recovery process undertaken by South Korean law enforcement authorities illustrates the challenges inherent in cross-border investigations of cryptocurrency-related crimes. After determining that a portion of the stolen cryptocurrency had been converted into bitcoin and held at a Switzerland-domiciled cryptocurrency exchange, the SKPA submitted compelling evidence to Swiss prosecutorial authorities demonstrating that these assets were linked to the theft from Upbit. What followed was an extensive **multiyear campaign** of asset recovery efforts. The SKPA encompassed numerous virtual and telephonic consultations alongside direct engagements at **the Swiss Federal Prosecutor's Office**. Working in concert with the South Korean Supreme Prosecutors' Office and the Ministry of Justice, the SKPA maintained persistent cross-border legal cooperation with Switzerland for nearly four years. Unfortunately, the foreign cryptocurrency exchanges **largely declined** to respond to South Korean law enforcement entities' overtures seeking the return of the illicitly obtained assets. This case demonstrates that close and timely cooperation with foreign virtual asset service providers and national financial intelligence units is imperative for effectively addressing cryptocurrency theft and recovering stolen crypto proceeds.

FACILITATORS OPERATING IN THIRD COUNTRIES

There are facilitators in third countries that contribute to money laundering and cash-out by DPRK actors. Mostly, these actors rely substantially on Chinese-based facilitators. Some Chinese facilitators collaborated with North Korean actors to procure fake identifications and operate a peer-to-peer cryptocurrency trader. Others even worked as **brokers** for money laundering and cashing out by purchasing crypto assets from DPRK actors and delivering converted funds in fiat currency in exchange for a certain amount of money in crypto or fiat form. In 2020, two Chinese nationals were **indicted** by the U.S. Department of Justice for laundering North Korean-linked cryptocurrency. In 2023, Sim Hyon-Sop, working for North Korean Foreign Trade Bank, was **charged** by the Department of Justice in three different indictments for his money laundering activities. He allegedly received cryptocurrencies stolen through hacking, remitted them to multiple wallets, and handed them to **brokers he contacted** in China or the United Arab Emirates. Without cooperation from states where facilitators operate, however, the justice process cannot proceed, and the charged individuals have never been arrested or tried. Information on the facilitators based in China, including ones involved in cryptocurrency laundering, has been delivered to China by **the participating states in the MSMT**. Unfortunately, there has been no indication that China took any action in response to such information.

Financial infrastructures in Southeast Asian countries often serve as a critical conduit for North Korean cryptocurrency laundering. For instance, between 2022 and 2024, a North Korean national with connections to Reconnaissance General Bureau **regularly met** in person with officials from Huione Pay, a Cambodia-based payment service, to convert cryptocurrency into fiat currency and execute international transfers. Further, North Korean actors based in Cambodia allegedly **used accounts** registered at Huione Pay in 2023 to move funds from the Ronin Bridge theft. No criminal indictment in relation to Huione Pay has yet been filed. Instead, the U.S. Treasury Department **excluded** Huione Group from the U.S. financial system. Several states **expressed concerns** to the Cambodian government regarding Huione Pay's activities facilitating North Korean actors' money laundering. Subsequently, the National Bank of Cambodia **revoked** Huione Pay's payments license; notwithstanding this action, the company seems to **continue its operations** within Cambodia.

SOUTH KOREA'S EFFORTS IN THE CHASE

South Korean crypto exchanges have been hit by North Korean malicious activities in cyberspace. In 2017, Youbit was hacked **two times** and entered insolvency. In addition to **the 2019 attack**, Upbit also had Solana-based crypto assets valued at **KRW 44.5 billion** stolen in 2025. In order to combat crypto theft by DPRK actors, South Korea has taken key steps in the right direction, which include introducing or reforming several crypto-related policies—including amending the Act on Reporting and Using Specified Financial Transaction Information (**SFTI Act**) and passing into law the Act on Virtual Asset Users Protection (**VAUP Act**)—and deepening international cooperation with the United States, Japan, and the FATF.

First, information sharing among virtual asset service providers (VASP), Korea Financial Intelligence Unit (KoFIU), and investigative entities has improved. On April 19, 2022, dedicated **direct communication lines** were established between the SKPA and five major domestic virtual

asset exchanges—Gopax, Bithumb, Upbit, Korbit, and Coinone. Later that year, on October 13, the SKPA concluded an **MOU** with these exchanges to facilitate investigative cooperation in virtual asset-related crimes and to enhance measures for victim protection and harm prevention. The information-sharing obligations of VASP and the KoFIU were also strengthened by amending the SFTI Act. In accordance with this act, VASP must report transactions suspected of involving illicit assets to the KoFIU (SFTI Act, Article 4). The KoFIU must provide information regarding specified financial transactions to investigative, financial, or intelligence entities—including the prosecutor general, director of the National Intelligence Service, and other relevant entities—where deemed necessary for including criminal investigations related to illicit proceeds or money laundering (SFTI Act, Article10). In accordance with the principle of reciprocity, the KoFIU may provide information on specified financial transactions to foreign financial intelligence units and receive such information in return (SFTI Act, Article 11).

Second, crypto exchanges' authority and responsibility in managing virtual assets under their custody have been expanded. Exchanges have been vested with discretionary authority to block virtual asset deposits and withdrawals without users' consent (VAUP Act, Article 11). Under this authority, exchanges can suspend deposits and withdrawals without users' consent where there are reasonable grounds to suspect involvement in money laundering activities. Exchanges are also required to hold no less than 80 percent of the economic value of users' virtual assets **isolated from the internet for security**, including in cold wallets (Decree to the VAUP Act, Article 11; Financial Service Commission Notice on the Supervision of Virtual Asset Business, Article 9).

Third, South Korea has worked together with the United States to counter North Korea's crypto theft. In 2022, South Korea established **a joint working group** with the United States to comprehensively address and respond to North Korea's lucrative activities in the cyber domain. This working group is composed of multiple agencies from the foreign affairs, intelligence, finance, and justice fields. In 2024, South Korea and the United States jointly imposed **sanctions** on the same person, Sim Hyon-Sop, for his money laundering activities. In 2025, South Korea, the United States, and Japan issued **a joint statement** on DPRK cryptocurrency thefts, where the three nations specifically emphasized the collaboration among the public and private sectors. Further, FATF members, including South Korea, have maintained North Korea's designation as **a high-risk jurisdiction** and called on all jurisdictions to cut off or restrict financial relationships with North Korea, implement enhanced financial monitoring, and strictly enforce UN sanctions.

Seeking International Law Enforcement Framework for Countering Crypto Theft

As illustrated above, South Korea has realigned its domestic policies to bring the virtual asset financial system within the scope of existing anti-money laundering/countering the financing of terrorism (AML/CFT) frameworks, and has deepened international cooperation with the United States, Japan, and the FATF to counter the North Korea's cryptocurrency theft. Nevertheless, absent effective cooperation at the global level—particularly from jurisdictions that enable North Korea's money-laundering activities—North Korea's pursuit of cryptocurrency will remain largely

unconstrained. Accordingly, there is a pressing need for an international cooperation framework that enables states to collaborate swiftly in identifying, freezing, seizing, and recovering stolen cryptocurrency, as well as in investigating and prosecuting cyber-enabled cryptocurrency theft. This section assesses the extent to which existing multilateral instruments—including the UN Convention against Transnational Organized Crime (UNTOC), the UN Convention against Corruption (UNCAC), the FATF recommendations, and the newly adopted UN Convention on Cybercrime—provide an adequate legal basis for such cooperation.

ASSESSING EXISTING FRAMEWORKS

The Budapest Convention on Cybercrime does not regulate money laundering. **UNTOC** and **UNCAC** do so, but are not fully suitable to regulate cyber-enabled crypto theft. The FATF provides **practical guidance** for states to invoke when requesting mutual legal assistance in investigations and prosecutions related to crypto theft and laundering. However, this guidance is not legally mandated, and thus implementation remains voluntary for each state.

UNTOC AND UNCAC

The UNTOC and the UNCAC include provisions requiring parties to impose customer identification, record-keeping, and suspicious transaction reporting obligations on financial institutions (UNTOC Article 7[1][a]; UNCAC Article 14[1][a]). The treaties further oblige parties to ensure that competent authorities responsible for combating money laundering—including administrative, regulatory, law enforcement, and, where appropriate, judicial bodies—are legally and institutionally empowered to cooperate and exchange information at both the domestic and international levels, subject to the conditions prescribed by domestic law (UNTOC Article 7[1][b]; UNCAC Article 14[1][b]). In addition, both treaties provide mechanisms enabling interstate cooperation for the identification and confiscation of criminal proceeds (UNTOC Article 13; UNCAC Article 55). The UNTOC has 195 parties, including South Korea, North Korea, the United States, and China, while the UNCAC has 192 parties, including South Korea, the United States, and China. Despite their near-universal participation, however, neither treaty offers an effective legal basis for seeking interstate cooperation specifically in cases of DPRK cryptocurrency theft and laundering, given their limited material scopes—the UNTOC to organized crime, and the UNCAC to corruption-related offenses.

FATF RECOMMENDATIONS

As highlighted above, timely cooperation from foreign financial institutions (i.e., cryptocurrency exchanges), is key for law enforcement authorities to identify, freeze, seize, or recover stolen cryptocurrency in their custody. For the purpose of law enforcement, cooperation from cryptocurrency exchanges can be sought based on **international AML/CFT standards** developed by the FATF. The FATF is an intergovernmental body established in 1989, **the objectives** of which are to protect financial systems from money laundering and the financing of terrorism and proliferation. Thirty-eight countries including South Korea, the United States, the United Kingdom, China, and Japan are FATF members. FATF has upgraded its international standards on AML/CFT regularly, in particular, to make virtual assets subject to the standards. The latest update's Recommendation No. 15 **makes clear** that "countries should ensure that virtual asset service providers are regulated for AML/CFT purposes."

FATF's recommendations associated with law enforcement provide practical guidance where countries can offer mutual legal assistance to each other in relation to crypto laundering. The FATF recommended that countries should ensure that investigative powers and techniques relating to the production, search, and seizure of information from financial institutions including crypto exchanges are available for use in response to requests from mutual legal assistance (Recommendation No. 37). In addition, countries should have measures to take expeditious action in response to requests by foreign countries seeking assistance to identify, trace, evaluate, investigate, freeze, seize, and confiscate criminal property, which includes recognizing and enforcing foreign orders on freezing, seizing, or confiscating (Recommendation No. 38). While the recommendations adopted by FATF are a practically valuable instrument on which law enforcement entities of each country can rely for cooperation, it is not legally binding.

FRAMEWORK UNDER THE UN CONVENTION AGAINST CYBERCRIME

Ultimately, the UNCC is the only internationally binding instrument to date that provides substantive and procedural rules for global cooperation to combat crypto theft and laundering by DPRK actors. Interestingly, both the DPRK, the primary perpetrator of cyber-enabled cryptocurrency theft, and China, the principal jurisdiction providing a safe haven for DPRK-affiliated actors, have signed the UNCC. The UNCC, adopted by the UN General Assembly in December 2024, remains unsigned by the United States and its like-minded countries, including South Korea, due to a range of unresolved controversies. One major concern is that the procedural provisions governing access to electronic data are perceived to lack sufficient human rights safeguards, thereby creating the risk that authoritarian states (including China, Russia, and North Korea) could exploit these mechanisms for surveillance purposes. Nevertheless, the UNCC does contain substantive and procedural rules that could contribute to seeking cooperation from foreign crypto exchanges (albeit indirectly) and countries that facilitate North Korea's crypto theft and subsequent laundering processes. In this context, the value that the convention adds for countering North Korea must be assessed.

ELIMINATING SAFE HAVENS BY CRIMINALIZING CRYPTO LAUNDERING

As illustrated above, DPRK actors' laundering of stolen cryptocurrency and cashing out is facilitated by third-party laundering services and enablers located in third countries. Thus, there is a need for globally harmonized rules where activities related to cryptocurrency laundering are criminalized and punishable. The UNCC requires parties to criminalize money laundering of proceeds of cybercrimes (Article 17) ("Cybercrimes" here refer to the offenses established under Articles 7–16 of the convention). Under the convention, proceeds of crime refer to "**any property**" derived from or obtained, directly or indirectly, through the commission of an offense. This convention makes clear that "property" means "assets of every kind," and includes "virtual assets" (Article 2[i]). Based on these provisions, parties are obligated to criminalize acts of laundering cryptocurrency stolen through hacking.

This criminalization obligation is linked to obligations for international cooperation. For instance, in accordance with the convention, parties must cooperate with each other for the purpose of sharing e-evidence relating to the offences established under the UNCC, namely, not only cyber-enabled criminal activities, but also crypto laundering (Article 35[1][b]). Further, parties must take effective

measures to cooperate with each other in conducting inquiries concerning the movement of criminal proceeds (Article 47[1][b]). In sum, these clauses can function as a legitimate basis to ensure that countries providing safe havens to facilitators assisting North Korea's crypto laundering activities can no longer do so. Instead, they are legally obligated to cooperate with other treaty parties in sharing e-evidence related to cryptocurrency laundering and conducting inquiries into the movement of stolen cryptocurrency.

INTERNATIONAL COOPERATION FRAMEWORK FOR IDENTIFYING, FREEZING, SEIZING, AND RECOVERING STOLEN ASSETS

The UNCC contains rules on international cooperation for the purpose of confiscating crypto proceeds generated through cybercrime. Parties first must adopt a domestic legal framework where criminal proceeds can be identified, traced, frozen, seized, and confiscated. In other words, parties are required to implement measures within their legal frameworks to confiscate criminal profits as well as equipment involved in the offenses (Article 31[1]). To facilitate confiscation, they must also adopt measures to enable authorities to identify, trace, freeze, or temporarily seize such items (Article 31[2]).

Second, The UNCC sets forth procedures through which parties may request the confiscation of criminal proceeds in cryptocurrency located in other parties. When proceeds of crime are situated in one party state, another party state that has jurisdiction over the predicate offenses can request the confiscation of the proceeds from it. In such a request, the requesting country must submit information on a description of the property to be confiscated, the location, and a statement of fact to the extent possible. Subsequently, after this request, the requested country is obligated to take measures to identify, trace, and freeze or seize proceeds of crime for such confiscation (Article 50).

Third, there is a provision where investigative entities could indirectly gain cooperation from foreign crypto exchanges. Each party must empower its courts or other competent authorities to order that bank, financial, or commercial records be made available or be seized for the purpose of international cooperation for confiscation (Article 31[7]). Based on this clause, a party can gain information regarding allegedly stolen cryptocurrency in the custody of foreign exchanges. Given that the SKPA encountered the foreign cryptocurrency exchanges' **refusal** to respond to the agency's requests in order to seek the return of the assets unlawfully obtained from the 2019 Upbit hack, these provisions may provide a legal ground on which a party can rely to draw cooperation from foreign crypto exchanges.

IN SUM

This assessment of existing international frameworks shows that the UNTOC and the UNCAC offer useful legal tools, but their limited scope prevents them from adequately addressing the specific challenges posed by cyber-enabled asset theft. The FATF recommendations similarly provide practical guidance for states, yet their nonbinding character means that implementation ultimately depends on each country's willingness to comply. As a result, even well-developed domestic regulations and strong bilateral partnerships are insufficient to fully address a threat that exploits gaps between national jurisdictions. In light of these limitations, the UNCC marks an important shift toward a binding global framework. By requiring the criminalization of cryptocurrency

laundering and establishing coordinated procedures for asset recovery, the convention advances international cooperation.

Conclusion

The decentralized and pseudonymous nature of cryptocurrency transactions provides North Korea with incentives to target crypto assets. This structural advantage is further compounded by sophisticated obfuscation techniques—including mixing, swapping, and cross-chain bridging—and anonymity-enhanced cryptocurrencies, all of which significantly reduce or eliminate transactional traceability. In response, law enforcement agencies have targeted cash-out points on CEXs that criminal actors exploit primarily to convert crypto assets into fiat currency, with a view to tracing, freezing, seizing, and confiscating stolen cryptocurrency. However, the absence of a global framework through which law enforcement authorities can seek cooperation from foreign cryptocurrency exchanges makes it difficult for the authorities to effectively and swiftly address cryptocurrency theft and recover stolen assets. Furthermore, facilitators operating within third countries are providing assistance in the money laundering process involving cryptocurrency stolen by North Korea. Unless law enforcement action is taken against those involved in money laundering who operate in safe havens, it will be difficult to prevent North Korea from laundering funds illicitly obtained through cyber activities.

From this perspective, law enforcement responses to North Korea's cyber-enabled theft of cryptocurrency and subsequent money laundering require global cooperation. In particular, states should seek harmonization at the level of interstate criminalization by defining cybercrime and money laundering of proceeds generated by cybercrime as criminal offenses. In addition, a framework is needed to ensure that law enforcement cooperation in identifying, freezing, seizing, and recovering stolen funds can be carried out swiftly and effectively. While the UNCC remains a subject of intense debate regarding human rights and surveillance, it nonetheless presents a significant opportunity to anchor a multilateral law enforcement regime. By mandating the criminalization of cryptocurrency-based money laundering and institutionalizing international cooperation for asset recovery, the convention could provide a more coherent and binding legal framework to address DPRK crypto theft and money laundering practices.

Joohui Park *is a senior researcher on the Cybersecurity Policy Research Team at the National Security Research Institute in the Republic of Korea.*

Responding to the Evolution and Global Expansion of the DPRK IT Worker Threat

By Yena Kim and Donghee Kim

The recent large-scale **indictment** of Democratic People's Republic of Korea (DPRK) information technology (IT) workers by the U.S. Department of Justice clearly demonstrates the escalating sophistication and severity of North Korea's cyber operations. This law enforcement action underscores not merely the detection of illegal activity, but the fact that the DPRK is leveraging cyberspace in a systematic manner to generate revenue for the regime. As of 2024, DPRK IT workers are estimated to generate between **$350 million and $800 million annually**.

As UN sanctions have constricted traditional sources of foreign currency, the DPRK has increasingly relied on overseas IT workers disguised as legitimate employees. What initially appeared to involve ordinary software development or application design has evolved into a coordinated strategy of infiltration, revenue generation, and, in some cases, data theft and financial crime.

This report examines the evolution of DPRK IT workers' operational tactics, the international expansion of their networks, and the limitations of existing responses. It argues that advisory- and sanctions-based measures alone are insufficient and calls for more proactive, government-led coordination and outreach to address this growing threat.

The Evolution of DPRK IT Workers' Operational Tactics

Since around 2017, DPRK IT workers have been creating fraudulent profiles, forged identities, and fake resumes to infiltrate freelance platforms and secure employment abroad. Initially, their work appeared to involve ordinary software development or application design, delaying recognition of the threat and enabling rapid proliferation.

During the hiring process, when video interviews were requested, accomplices appeared on camera while the DPRK worker redirected the conversation to a phone or chat-based format, citing supposed technical issues. In some cases, they remotely accessed the accomplice's device to perform technical demonstrations during the interview.

Once remote work feasibility was confirmed, DPRK IT units collaborated with U.S.-based conspirators to operate "laptop farms," enabling multiple simultaneous infiltrations. U.S. accomplices received and managed employer-issued laptops, while DPRK operatives remotely controlled them to impersonate numerous employees across several companies.

These workers do not merely conduct software development. They strategically infiltrate defense contractors, government contractors, and critical infrastructure organizations to obtain sensitive internal access. Once inside, they exfiltrate confidential data and intellectual property to DPRK or third-party criminal networks. Some groups install backdoors for follow-on operations or threaten data leaks after termination to extort money. Multiple cases show DPRK IT workers securing developer positions at cryptocurrency companies and directly stealing hundreds of millions of dollars in assets.

More recently, DPRK IT workers have aggressively integrated artificial intelligence (AI) into their disguise and infiltration tactics. They generate fake profile images, automatically produce large volumes of resumes and cover letters using generative AI tools, manipulate social media activity, and use deepfake video and audio during remote interviews.

These AI-enabled methods represent a shift from basic identity fraud to a highly structured and coordinated global infiltration strategy. This trend reinforces the need not only for companies, but also governments, to strengthen both defensive and policy responses, as these operations increasingly constitute a direct national security risk.

The International Expansion of DPRK IT Workers and Accomplices

DPRK IT workers and their accomplice networks have expanded their geographic scope beyond China and Russia—where they had traditionally operated—to new third-country destinations. Since 2020, they have increasingly been deployed to **Equatorial Guinea, Laos, and the United Arab Emirates**, and more recently to **Cambodia, Guinea, Nigeria**, and **Tanzania**, where they conduct remote work for U.S. companies to earn foreign currency. However, in the second half of 2024, as U.S. law enforcement measures intensified, some workers **shifted** operations to Europe, where oversight remains relatively less stringent.

In Europe, IT workers have engaged in coordinated activities across multiple countries—including Germany, Portugal, and the United Kingdom—leveraging numerous personas and AI-generated synthetic profiles to register on local employment platforms and global freelancing websites, and to apply for positions at defense contractors and government-related organizations. One notable case was identified in Serbia, where a DPRK IT worker infiltrated a token corporation and stole approximately **$175,000 worth of cryptocurrency**, subsequently transferring the funds to overseas accounts. This case demonstrates that North Korea is actively exploiting European financial and crypto-asset systems to construct a foreign currency procurement structure.

Such activities have also been detected across Asia. In Japan, multiple incidents have been uncovered involving DPRK IT workers engaged in fraudulent employment and illicit operations. In China, a North Korean worker attempting to **steal** military technology was arrested and detained, illustrating that the activities of these workers are expanding beyond conventional software development into strategically sensitive areas such as defense technology. These incidents indicate that their operations are no longer confined to North America and Europe; instead, they are spreading across Asia and expanding globally in both geographic and industrial dimensions.

These overseas operations would not be possible without the support of accomplices operating worldwide. **Accomplice networks** in the United States, Japan, Ukraine, and the United Arab Emirates collaborate with DPRK IT workers by managing laptop farms, laundering funds, managing stolen assets, forging identities, and establishing illicit remittance channels. This support network forms the critical infrastructure that enables the covert operations, revenue generation, and sanctions evasion activities of DPRK IT workers.

As North Korea's concealment techniques and operational sophistication continue to advance, and as its networks of workers and accomplices expand globally, it is increasingly clear that the current international cooperation framework alone has limitations in mounting an effective response.

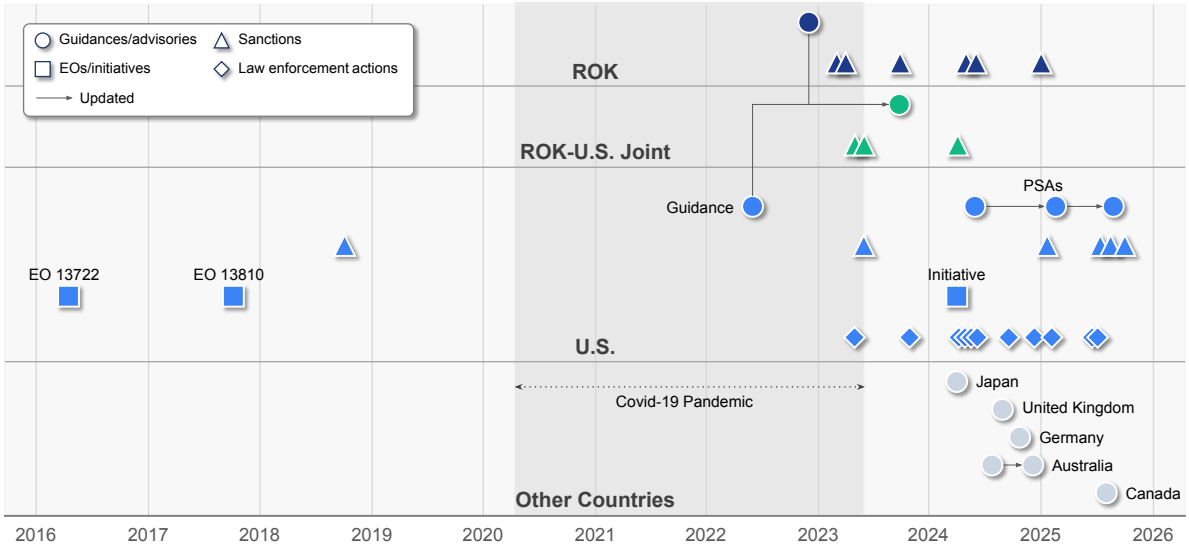
Limitations of International Responses to DPRK IT Activities

The United States is the primary country where DPRK IT workers are most actively operating, and it has responded with the most robust measures, including various sanctions, to curb their employment activities within the country. As a result, by the second half of 2024, DPRK IT workers shifted their operations primarily to Europe, where oversight is relatively less stringent. Figure 1 summarizes guidance and advisories, sanctions, executive orders (EOs) and initiatives, and law enforcement actions issued by the Republic of Korea (ROK), the United States, and other countries, organized by release date. The figure illustrates that DPRK IT workers have long been recognized by the U.S. government as a serious threat, and the United States has actively responded through EOs, initiative, sanctions, and legal actions.

Initially, EOs **13722** and **13810** designated DPRK IT workers' overseas deployment and related activities as sanctions targets, prohibiting their provision of software and technical services. However, the **Covid-19 pandemic** (March 2020–May 2023) led to a surge in remote work, enabling DPRK IT workers to work for U.S. companies remotely. Consequently, since 2022, U.S. government and intelligence

agencies have issued a total of four documents, including one **guidance** and three **public service announcements** to alert companies to risk indicators and response measures. From 2023 onward, interagency collaboration among the U.S. Departments of Treasury, Justice, and State has intensified, resulting in large-scale sanctions and prosecutions targeting DPRK IT workers and networks.

Figure 1: International Responses to DPRK IT Workers



Source: Government advisories, sanctions announcements, and official statements from the Republic of Korea, the United States, Japan, the United Kingdom, Germany, Australia, and Canada.

Table 1: International Responses to DPRK IT Workers by Year

Year	Country	Response Type	Date
2016	United States	Executive Order 13722	March 15, 2016
2017	United States	Executive Order 13810	September 20, 2017
2018	United States	Sanctions	September 13, 2018
2022	United States	Guidance	May 16, 2022
2022	ROK	Advisory	December 8, 2022
2023	ROK	Sanctions	February 10, 2023; March 21, 2023; September 1, 2023
2023	ROK–United States Joint	Coordinated sanctions	April 24, 2023; May 23, 2023

2023	ROK–United States Joint	Joint advisory	October 19, 2023
2023	United States	Sanctions	May 23, 2023
2023	United States	Law enforcement actions	April 24, 2023; October 18, 2023
2024	ROK	Sanctions	April 3, 2024; May 24, 2024; December 26, 2024
2024	ROK–United States Joint	Coordinated sanctions	March 28, 2024
2024	United States	Public service announcement	May 16, 2024
2024	United States	Initiative	March 2024
2024	United States	Law enforcement actions	May 15, 2024; May 15, 2024; May 16, 2024; May 16, 2024; August 8, 2024; December 12, 2024
2024	Japan	Advisory	March 26, 2024
2024	Australia	Advisory	August 26, 2024; December 14, 2024
2024	United Kingdom	Advisory	September 12, 2024
2024	Germany	Advisory	October 1, 2024
2025	United States	Public service announcements	January 23, 2025; July 2, 2025
2025	United States	Sanctions	January 16, 2025; July 8, 2025; July 24, 2025; August 27, 2025
2025	United States	Law enforcement actions	January 23, 2025; June 5, 2025; June 30, 2025
2025	Canada	Advisory	July 16, 2025

Source: Government advisories, sanctions announcements, and official statements from the Republic of Korea, the United States, Japan, the United Kingdom, Germany, Australia, and Canada.

In March 2024, the United States launched the "**DPRK RevGen: Domestic Enabler Initiative**," prioritizing the identification and shutdown of laptop farms. The initiative disclosed multiple law enforcement actions against accomplices hosting laptop farms, DPRK IT workers involved, domains used for illicit employment, and associated revenues. Particularly, from the second half of 2025,

large-scale joint law enforcement actions and multiple sanctions targeting DPRK IT workers have been publicly disclosed, reflecting a significant strengthening of U.S. response efforts.

Following the United States, the ROK has been the most active country in responding to DPRK IT worker operations. In December 2022, the ROK issued its first **joint government advisory** calling for enhanced verification of DPRK IT worker identities. Since then, the ROK has independently designated six entities and individuals associated with DPRK IT workers involved in overseas foreign currency-earning activities as sanctions targets. The ROK and the United States have also coordinated in responding to DPRK IT workers. In April 2023, the two countries jointly designated targets for sanctions for the first time, and to date, there have been three instances of joint target designations. Separately, in October 2023, **both countries** issued a joint government advisory updating their respective advisories with the latest trends.

In 2024, countries including **Australia**, **Canada**, **Germany**, **Japan**, and the **United Kingdom** also joined the international effort by issuing advisories on DPRK IT workers' disguised employment. This development reflects a growing recognition that the DPRK's covert employment activities are not merely a foreign currency-earning issue, but also an international security and economic risk. Overall, Figure 1 demonstrates that the principal actors responding to DPRK IT worker operations have expanded from the ROK and the United States to other countries, including those in Europe, highlighting the need for a broader international response framework.

Figure 1 also highlights certain challenges. According to U.S. indictments, DPRK IT workers have been active since at least **2017**. Their initial activities were disguised as legitimate employment, which delayed recognition of the associated risks by both the international community and private companies. The figure shows that substantial international responses only began in earnest around 2024.

During the Covid-19 pandemic, as remote work culture expanded, DPRK IT workers found it increasingly easy to participate in overseas corporate projects by posing as third-country workers. During this period, they are also believed to have rapidly enhanced both their technical capabilities and their identity-masking and money-laundering techniques. At the same time, international sanctions enforcement and broader multilateral responses remained largely stagnant, creating a widening gap between North Korea's evolving methods and the international community's ability to counter them.

As a result, advisory and sanctions-based responses have inherent limitations. DPRK IT workers are state-sponsored and systematically managed, allowing them to continuously develop innovative techniques and methods. In practice, new identity-masking and technical methods have repeatedly emerged immediately following sanctions. While the United States has combined executive orders, sanctions, prosecutions, and law enforcement actions, these measures alone have had limited effectiveness, as North Korea consistently develops new evasion methods to circumvent sanctions.

Consequently, advisory- and sanctions-based responses alone remain insufficient to fundamentally disrupt DPRK IT worker operations. While detection and response at the company level are

important, the burden of identifying and mitigating these threats has largely been placed on individual companies, many of which lack the resources, visibility, or expertise to do so effectively. Moreover, the transnational and adaptive nature of DPRK IT worker networks limits the effectiveness of isolated, reactive measures.

These limitations underscore the need for a more proactive and coordinated approach led by governments. In particular, systematic outreach to domestic companies, structured engagement with international partners, and sustained public-private collaboration are essential to closing existing response gaps. Without stronger government-led efforts to disseminate risk awareness, share intelligence, and align international countermeasures, existing responses will continue to lag behind the evolving tactics of DPRK IT workers.

Government-led Outreach and Coordination to Counter DPRK IT Worker Threats

DPRK IT worker operations are expected to persist through 2026 and continue expanding geographically as enforcement pressure intensifies in the United States. These operations are likely to leverage increasingly sophisticated techniques, including multimodal generative AI—including voice, text, and video deepfakes—to sustain disguised employment. At the same time, cryptocurrency-related companies and other digitally native sectors will remain **high-value targets** due to their remote work structures and the potential for rapid financial gain.

International sanctions and law enforcement actions have imposed tangible costs on these networks, but they have not fundamentally disrupted DPRK IT worker operations. The adaptive, state-sponsored nature of these activities enables rapid evolution in response to enforcement measures. Reliance on reactive, company-level detection alone places an unsustainable burden on individual companies and creates persistent gaps between emerging threats and effective countermeasures. Addressing these limitations requires a shift toward proactive, government-led outreach and coordination.

LESSONS FROM THE PRIVATE SECTOR: IMPLICATIONS FOR GOVERNMENT ACTION

Companies face significant structural constraints when disclosing DPRK IT worker incidents. When a company inadvertently hires a DPRK IT worker, the resulting risks extend beyond technical compromise to reputational damage, customer attrition, legal exposure, and contractual liabilities. These factors strongly discourage voluntary disclosure, even when early reporting would benefit the broader ecosystem.

Nevertheless, in 2024, one cybersecurity company publicly **disclosed** an attempted internal intrusion linked to DPRK IT worker activity. The company documented the full recruitment process—including job postings, interview procedures, and identity verification steps—and immediately shared all relevant data with U.S. cybersecurity firm Mandiant and the FBI to support early-stage investigations. The company also publicly release lessons learned and response measures through its website.

Following the incident, the company **conducted** organization-wide employee training, implemented fingerprint-based identity verification, and restricted corporate laptop delivery to verified UPS shipments requiring photo identification. In addition, the company published a comprehensive **white paper** detailing insider-threat risks and preventive controls, and it continued to release updates on DPRK IT worker activity trends.

According to the company's CEO, the decision to disclose was intended to raise awareness of the widespread nature of DPRK IT worker infiltration attempts and to warn other organizations of comparable risks. By transparently sharing its experience, the company contributed to elevating industry-wide security standards while reinforcing customer trust.

However, such best practices remain fragmented across jurisdictions and sectors, making comprehensive access difficult for most companies—particularly small- and medium-sized enterprises. This fragmentation highlights the limitations of relying solely on voluntary private sector disclosure. Accordingly, governments should assume a more active role in collecting, standardizing, and disseminating these lessons through sustained outreach efforts.

CENTRALIZED INFORMATION SHARING AND INSTITUTIONAL VERIFICATION FRAMEWORKS

Governments already possess extensive intelligence related to DPRK cyber operations and IT worker networks, yet this information often fails to reach companies in a usable and timely manner. To close this gap, the ROK government should centrally aggregate and disseminate both best practices and DPRK IT worker–specific indicators—including email addresses, recurring account naming patterns, profile photos, commonly cited education and career information, and IP addresses associated with laptop farms.

Rather than creating new mechanisms, existing government–private sector information-sharing platforms focused on DPRK cyber threats should be expanded to incorporate these indicators. Centralization would significantly lower access barriers, enabling companies of all sizes to integrate this information directly into recruitment screening, identity verification, and internal security training processes.

In parallel, the ROK should consider establishing an institutional employment-eligibility and identity verification framework analogous to the U.S. I-9 and E-Verify systems. In the United States, these mechanisms have been used not only to confirm work authorization, but also to flag identity inconsistencies and third-country impersonation patterns relevant to DPRK IT worker investigations. A comparable system in the ROK would provide companies with an additional safeguard when pre-employment screening fails.

REPORTING, INCENTIVES, AND ENFORCEMENT MECHANISMS

Timely reporting remains critical to preventing escalation and secondary harm once DPRK IT worker activity is detected. However, incentives alone are unlikely to overcome companies' reluctance to disclose incidents. A dual-track framework combining legal protections and meaningful penalties is therefore required to make reporting the rational choice rather than concealment.

The government should provide whistleblower protections and limited liability safeguards for companies that report suspicious activity promptly. At the same time, penalties—including increased fines—should be strengthened for companies that knowingly or negligently employ DPRK IT workers or fail to report suspicious indicators within a defined time frame. Enforcement should also extend to operational enablers, such as laptop farm operators and intermediaries who facilitate overseas employment and financial flows, as these actors are essential to sustaining DPRK IT worker networks.

INTERNATIONAL OUTREACH AND THE ROK'S STRATEGIC LEADERSHIP

Although remote work is less prevalent in the ROK than in many Western economies, the country faces unparalleled national security exposure due to its geographic proximity to North Korea. This position confers both heightened risk and a unique responsibility to lead international awareness and coordination efforts.

Recent reporting indicates that DPRK IT worker activity has increased across Europe, where remote work adoption remains high and threat awareness comparatively low. The ROK should proactively engage European partners to emphasize that weak awareness in remote work environments can escalate into broader national and economic security risks. Similar outreach is warranted in Southeast Asia, Africa, and other regions where awareness remains limited.

In parallel, bilateral cooperation between the ROK and the United States—as well as trilateral coordination with Japan—should continue through regular advisories, working-level exchanges, and joint statements. Given the increasingly blurred line between DPRK IT workers and overtly malicious cyber actors, a new joint advisory emphasizing the growing use of direct exfiltration and intrusion tactics would reinforce shared threat perception and strengthen collective deterrence.

Ultimately, countering DPRK IT worker operations requires more than cooperation—it requires leadership. By centralizing information, institutionalizing verification mechanisms, incentivizing reporting, and proactively engaging international partners, the ROK can move from a reactive participant to a global agenda-setter in addressing this evolving threat.

Yena Kim *is a senior researcher on the Cybersecurity Policy Research Team at the National Security Research Institute in the Republic of Korea.* ***Donghee Kim*** *is a senior researcher and manager of the Cybersecurity Policy Research Team at the National Security Research Institute in the Republic of Korea.*

Part IV

Recommendations for a Joint U.S.-ROK Cyber Resilience Strategy

Recommendations and Next Steps

South Korea's current approach to cybersecurity, defined by the 2024 National Cybersecurity Strategy and its subsequent Basic Plan, represents a fundamental shift from an ineffective and reactive defense to a more proactive posture. Its new policy is to attribute and impose costs on perpetrators. The most significant change in the 2024 strategy is the explicit commitment to offensive cyber defense. This is not about retaliation; it is about defense and disruption, specifically targeting North Korean cyber activities and infrastructure. The ROK has openly identified North Korea as the primary threat actor. The National Intelligence Service (NIS) and the ROK military have been given authority to identify and neutralize North Korean cyber networks and cyber threat infrastructure, such as command and control nodes, before cyberattacks are fully launched.

To operationalize the strategy, the government released the National Cybersecurity Basic Plan, with 100 actionable tasks (some of which are classified), including mandatory disclosures by listed companies and expanded legal authorities to track and seize cryptocurrency or other financial assets taken by North Korea. The tasks, intended to transition South Korea from a reactive defensive posture, and are organized into five strategic pillars:

- Bolstering Offensive Cyber Defense Operations
- Building Global Cooperation Mechanisms
- Enhancing Resilience of Critical Infrastructure
- Gaining a Competitive Edge in (cyber-related) Emerging Technologies
- Strengthening the ROK's Integrated Response Capability

The new strategy and the Basic Plan put the ROK at the leading edge of cybersecurity policy. The following recommendations build on actionable tasks to strengthen and operationalize these pillars:

- Further extend policy guidance and legal authorities. Building on the 2024 Cybersecurity Strategy, the ROK should develop detailed policy and refine legal frameworks on the use of disruptive countermeasures. This should also define threshold for "preemptive action." The goal is to embed cyber defense in a larger strategy to deter and constrain the DPRK.

  In some cases, extending legal authorities will require a careful balancing of the need for rapid action against essential protections for civil liberties by requiring a court-issued warrant before acting. On the whole, it is probably better to emphasize civil liberty protection over the requirements of active defense, but efforts to streamline the warrant process for the new policy are essential for its success.

  In addition to policy and legal frameworks, it may be useful to clarify governance arrangements and broader capacity building for active cyber defense, including closing gaps in planning, expanding legal frameworks, and strengthening operational capabilities—particularly within the military.

- **Streamline attribution requirements.** Attribution remains politically essential but has in the past become an obstacle to action. The ROK can adopt a streamlined attribution model for the DPRK. The DPRK's centralized governance system simplifies some attribution tasks and removes the need for "court-level" certainty before authorizing a disruptive response. The best way to do this is to adopt a "campaign-based" approach that looks at a consistent and continuous pattern of malicious activity rather than incident-level attribution.
- **Reinforce bilateral cooperation.** The ROK can improve cooperation with the United States by improving shared situational awareness, including by using an agreed Cyberattack Severity Classification Framework. Cooperation can also be expanded by collaborating on the disruption of cryptocurrency laundering and seizure of illicit funds, developing joint rules of engagement to allow for synchronized responses to malicious activities, and creating deconfliction mechanisms to avoid interference or fratricide. Expanded joint training and exercises can provide the operational experience needed for greater bilateral cooperation.

  It is also important to strengthen real-time (or near-real-time) information sharing and secure intelligence-sharing mechanisms. South Korea already has access to CISA's Automated Information Sharing system. This could be automated further through greater use of AI tools under the General Security of Military Information Agreement framework, by streamlining the United States' foreign disclosure process, expanding pre-approval processes, and adopting "Releasable to ROK" as a default classification.

- **Erode the DPRK's economic returns from malign cyber activity.** Use "whole-of-government" tools, including law enforcement and financial intelligence, to target the DPRK's financial pressure points and reduce the returns on cybercrime, particularly in the decentralized finance (DeFi) and cryptocurrency sectors. This can be done in coordination with multilateral efforts to reduce cryptocurrency crime. These efforts would

also be further strengthened by explicitly creating sustained operational friction for the DPRK, as well as financial disruption, such as disrupting IT worker activities, blocking fraudulent accounts, and dismantling botnets.

- **Create additional legal safeguards.** The ROK private sector should have additional mechanisms through which to share data with the ROK government, modeled after the U.S. Cybersecurity Information Sharing Act. Given South Korea's approach to regulation, it may be necessary to provide limited exceptions, such as "safe harbors," anonymization processes, and other legal protections to enable voluntary information sharing with the private sector. Measures like allowing NIS to issue temporary administrative orders for rapid action (within carefully prescribed limits) could be a useful addition.
- **Engage international stakeholders.** Creating a diplomatic strategy for cybersecurity is essential to the success of active cyber defense. A diplomatic strategy should expand coordination and cyber intelligence sharing with the "Five Eyes" partners and regional allies like Australia and Japan, and work with the European Union and multilateral enforcement bodies like Interpol and the Financial Action Task Force. It may also be worth expanding this to include cooperation with Southeast Asian countries, given their role as potential infrastructure and operational hubs for the DPRK.

  It is probably unrealistic to expect public cooperation with China regarding DPRK infrastructure on its soil. China has little incentive to cooperate. This would only change if DPRK activities imposed an unacceptable cost on China itself, something that is unlikely to occur. By allowing the use of its territory and infrastructure, China is, as recent UN reports have highlighted, a key enabler of DPRK criminal activity. Chinese support for the DPRK should be a potential agenda item for any bilateral engagements, but progress on DPRK-related crime will be difficult.

- **Strengthen strategic signaling.** The ROK should develop and use a public narrative to explain its cyber actions to the public, allies, and other nations while using indirect signaling to shape DPRK perceptions and manage any escalation risks. This could take the form of a statement from the Blue House or Ministry of Foreign Affairs. Joint messaging would amplify the security and diplomatic benefits and reduce any risks from active cyber defense. When appropriate, joint U.S.-ROK messaging, or messaging with Asian partners or the European Union, could further increase the impact of messaging.

  These steps are interconnected and form a nexus for active cyber defense. Better attribution, for example, can be reinforced by diplomatic efforts, particularly intelligence sharing with allies and information sharing with the private sector. Ultimately, these integrated efforts will transform the ROK's cybersecurity posture, increasing national resilience and international stability against the DPRK and other digital adversaries.

- **Develop a diplomatic strategy for international stakeholders.** A diplomatic strategy must be grounded in a recognition that South Korea must now play a larger role in the world to advance and safeguard its national interests. The ROK has global interests and its focus

cannot be solely on the Korean Peninsula anymore, particularly considering its expanding economic stature. Partners for controlling the DPRK's cyber activities must also be global. The ROK can expand coordinated diplomatic actions with the European Union, United Kingdom, regional allies like Australia and Japan, and Singapore and other Southeast Asian countries as part of a larger cyber diplomacy strategy.

While it would be a complicated diplomatic task, engagement with some of the other countries that allow DPRK cyber operations could help complicate DPRK cybercrime activities. These include several ASEAN member states and several states in the Middle East, such as the United Arab Emirates. ASEAN is probably the most likely avenue for successful engagement, starting with Singapore, and any diplomatic could be supported by the United States, Australia, and Japan.

About CSIS

The Center for Strategic and International Studies (CSIS) is a bipartisan, nonprofit policy research organization dedicated to advancing practical ideas to address the world's greatest challenges.

Founded in 1962, CSIS is led by General Joseph F. Dunford, who was appointed chief executive officer in 2026, succeeding John J. Hamre. The CSIS Board of Trustees is chaired by Thomas J. Pritzker, who has held the position since 2015.

CSIS brings together more than 275 full-time staff and a global network of affiliated scholars working across four core areas of public policy: defense and security, geopolitics and foreign policy, economic security and technology, and global development. Our scholars are regularly called upon by Congress, the executive branch, the media, and others to explain the day's events and offer recommendations to improve U.S. strategy.

CSIS's purpose is to define the future of national security. We are guided by a distinct set of values—non-partisanship, independent thought, innovative thinking, cross-disciplinary scholarship, integrity and professionalism, and talent development. CSIS's values work in concert toward the goal of making real-world impact.

What sets CSIS apart is not only the quality of its research but how that research is conveyed: through original data and open-source analysis; through multimedia and data visualization; and through connecting relevant analysis with key audiences at the moments that matter most. This combination of rigor, independence, and reach has made CSIS a driving force in the policy debates shaping American security and prosperity.

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

1616 Rhode Island Avenue NW, Washington, DC 20036
202.887.0200 | **www.csis.org**

FOLLOW CSIS