KGRI
Keio University Global Research Institute

KGRI Working Papers

No. 2

The Reality of Japan's Active Cyber Defense: Lessons Learned from the Evolution of Cyber Deterrence through U.S. and U.K. Cyber Operations Since the Late 2010s

Version 1.0

July 2026

Ichida Akira

Keio University Global Research Institute

Ph. D. student, Graduate School of Media and Governance, Keio University

The Reality of Japan's Active Cyber Defense:

Lessons Learned from the Evolution of Cyber Deterrence through U.S. and U.K. Cyber Operations Since the Late 2010s

Ichida Akira

Doctoral student at Keio University, Japan

Abstract

Cyberspace is referred to as the fifth operational domain, following land, sea, air, and space. With the rapid increase in the importance of cyberspace in modern society, it has become essential to deter cyberattacks and prevent further attacks. Under these circumstances, Japan finally enacted the “Cyber Response Capability Enhancement Acts” in May 2025, initiating active cyber defense. Meanwhile, cyber-advanced nations like the United States and the United Kingdom have already been conducting offensive cyber operations in the “the gray zone,” i.e., the area between armed conflict and peacetime, since the late 2010s. These operations are carried out not only against non-state actors, including hacker groups, but also against state actors, and information on counterattacks is disclosed through public attribution. The synergistic effects of these efforts have strengthened the cyber deterrence by punishment, which had previously been considered complementary. This study analyzed the results of cyber operations conducted by the U.S. and U.K. since the late 2010s to examine the effects and limitations of these cyber proactive operations on cyber deterrence by punishment and by denial thereby proposing reality and challenges for Japan.

Introduction

Cyberspace is referred to as the fifth operational domain,[1] alongside land, sea, air, and space. In modern society, critical infrastructure such as electricity, gas, and water supply relies on information and communication technology (ICT) that exists within this domain, and such critical infrastructure cannot function effectively without it. Consequently, cyberattacks are counted—together with extreme weather, natural disasters, large-scale illegal immigration, and terrorist attacks—as one of the five global risks.[2] Hence, deterring (i.e. preventing or intercepting) cyberattacks in cyberspace (hereinafter simply "cyber deterrence") is one of the key issues in national security. Nevertheless, looking at how cyber deterrence has evolved to date, it largely remains focused on passive denial-based deterrence measures that rely on cyber defense system such as firewall. Unlike nuclear deterrence, which rests on the premise of retaliatory strikes using nuclear weapons, the inherent difficulty in attributing an attack to a specific actor in cyberspace—so-called "attribution problem"—significantly constrains attempts to dissuade an adversary's intent to attack.

Meanwhile, comparing the policies and actual measures of the United States of America (U.S.) and the United Kingdom (U.K.) in the first half of the 2010s and from the latter half of the 2010s onward, one finds two clear shifts. First, to enhance deterrence by denial, there has been the evolution of "Active Cyber Defense," (hereinafter [ACD]), which entails proactively and early collecting, analyzing, and evaluating information beyond the system being defended, and taking measures accordingly. A notable example in the U.S. is the Hunt Forward Operations (hereinafter [HFO]), in other word an "overseas version of ACD," where the U.S. installs detection equipment abroad for intelligence gathering and other activities.

Second, having reached a certain level of resolution in the attribution problem, there have been cases suggesting that counterattacks were conducted under the framework of Offensive Cyber Operations (hereinafter [OCO]), which enable deterrence by punishment. These examples have begun appearing in news coverage. In particular, both the U.S. and the U.K. have disclosed that, in so-called the "gray zone" situations which do not rise to the level of armed conflict, they have carried out OCO abroad against nation-states or hacker collectives. The question arises: what significance do these operations, and their disclosure hold for cyber deterrence?

Meanwhile, in Japan, so-called Active Cyber Defense Acts (hereinafter Japan's ACD Acts) have finally been enacted. While the text represents a major step toward realizing Japan's ACD—almost equivalent to the U.S. and U.K.'s ACD and OCO—numerous challenges remain.

Therefore, to derive lessons for Japan, this paper sets the research question: "How can the operation of these proactive cyber operations by the U.S. and the U.K. enhance the effectiveness of cyber deterrence, from the gray zone to peacetime?" This is based on the initiatives being pursued by the U.S. and the U.K., particularly those aimed at improving cyber deterrence that has been underway since the late 2010s.

Indeed, upon investigating coverage of the U.S. and the U.K. capabilities to counter cyberattacks, one can see emerging evidence that might be considered to strengthen deterrence. This leads to the hypothesis: "ACD, HFO, and OCO under gray zone situations contribute to increasing both deterrence by denial and deterrence by punishment."

The analytical framework utilizes publicly available operational evidence from the latter half of the 2010s, with a particular focus on the U.S. Cyber Command (USCYBERCOM), established following the consolidation of U.S. military cyber operational capabilities, specifically its HFO and OCO. Additionally, the framework emphasizes the contributions of the U.K. National Cyber Security Centre (NCSC) and its ACD, as well as OCO conducted by the U.K. National Cyber Force (NCF). Specifically, after confirming in Chapter 1, based on prior research, what deterrence is and what constitutes deterrence power, we identify issues in cyber deterrence and ways to improve effectiveness. Chapter 2 examines how ACD, HFO, and OCO in gray zone can function as deterrence power in the context of international or domestic political developments in the U.S. and the U.K. In doing so, the analysis will be divided into three areas: ACD, HFO, and OCO in the gray zone. Building on the evolution of U.S. and U.K. cyber deterrence policy in the 2010s, it will examine the performance of specific operations—the U.S. HFO and OCO, and the U.K.'s ACD and OCO—from the late 2010s onward. The findings will be compared and verified against strategic documents to identify the potential impact these operations have on cyber deterrence. Chapter 3 provides a synopsis of the effects on both deterrence by denial and by punishment, in as well as their limitations. Based on the analysis of this research, it will show the necessary approaches to overcome anticipated capability concerns following the enforcement of Japan's ACD legislation. Chapter 4 compares the capabilities of ACD and OCO in the U.S. and the U.K.—and the cyber deterrence that results from their operational deployment—with the capabilities and deterrence expected in Japan following the enforcement of the ACD Acts. Based on these differences, the chapter identifies the measures necessary to address the anticipated challenges.

1. The Concept of Cyber Deterrence, Prior Research, Challenges, and Potential Improvements

(1) Cyber Deterrence

The term deterrence is defined by the U.S. Department of Defense as "the act of preventing an adversary from taking undesirable actions by sending tangible or intangible messages to the adversary, thereby causing the adversary to refrain from actions that it would otherwise have taken."[3] Joseph S. Nye, Jr. offers a definition of deterrence as "a psychological process that depends on the perceptions of both the actors and the targets, and the ability to communicate those views clearly" and "deterrence means dissuading someone from doing something by making them believe that the costs to them will exceed their expected benefit."[4] Aaron F. Brantly further asserts the need for "the mechanisms to establish that further costs are credibly wagered to provide clarity for an adversary."[5]

In other words, in a mechanism both entities trust, the defending side demonstrates its will and causes the attacker to abandon the assault. Broadly speaking, deterrence takes two forms: deterrence by punishment, which projects force to evoke the prospect of

retaliation and thus dissuade assault, and deterrence by denial, which compels the perpetrators to do a cost effectiveness calculation and abandon its attack. Collectively, these be termed as traditional deterrence. In comparing physical deterrence to cyber deterrence, it is often contrasted with nuclear deterrence from the U.S.–Soviet Cold War era. Nuclear deterrence is based around nuclear retaliation, i.e. deterrence by punishment, sustained by the existential deterrence power of nuclear weapons that could destroy an adversary multiple time over. Takahisa Kawaguchi notes that because nuclear weapons are so powerful, mere existence is sufficient to imbue them with deterrence power. However, referencing the definition from the U.S. Department of Defense, he observes that cyber deterrence relies less on existence and more on operational practice, such as continuously detecting and analyzing threats and vulnerabilities in real time and updating security systems. Kawaguchi emphasizes that "cyber intelligence, security systems upgrades, and other ongoing operations comprises the core of cyber deterrence."[6]

(2) Traditional Deterrence and Its Challenges

a. The Attribution Problem

The fundamental principle of deterrence by punishment is contingent upon the identification of the aggressors, their underlying motives, the objectives they seek to attain, and, in the event of a state actor being implicated, the attribution of responsibility to that state to facilitate a suitable retaliatory response.[7] However, in cyberspace, it remains challenging to definitively identify an attacker, a phenomenon referred to as the attribution problem. It is often the case that states decide against retaliation, in consideration of the risk of mistargeting, which engenders difficulties in the implementation of deterrence by punishment.

Furthermore, the inherent nature of cyberspace as an attack-dominant domain is a salient factor in this problem. It is impossible to eliminate the risk of intrusion entirely, and the cost implications of such an attack are negligible in comparison to the substantial investment required by the defender in systems, education, and other elements of cyber defense. For instance, there have been announced cases of 10 million lines of security code being breached by 125 lines of malware,[8] which vividly demonstrates that defending purely via deterrence by denial is, in terms of cost, unrealistic.

Concurrently, frameworks such as NATO's Tallinn Manual seek to establish norms and consequently a form of binding effect, and there are ongoing efforts to progressively develop customary international law. However, the stance of the U.S. and the U.K., regarded as responsible and democratic cyber powers espoused by which advocate for a free and open cyberspace, is in direct opposition to the stance of states such as China and Russia. These states contend that national sovereignty extends to domestic cyberspace.[9] Since the latter are improbable to retract their proposals, the mutual consent should not be anticipated.

b. Challenges within the Attribution Process

When a major cyber incident occurs, the attribution process at the national level provides guidance on how the state should respond. In 2015, Thomas Rid and Ben Buchanan described "attribution is a nuanced and multi-layered process not a problem that can

simply be solved and not be solved," and "Governments get to decide how to do attribution, and they get to decide when attribution is good enough for action."[10]

Figure 1 is a diagram of the OODA loop[11] for attribution by the state in the event of a major cyber incident. It consists of the following processes: Detection and Reporting (Observe: either the public or private sector detects the attack and reports it to the relevant authorities), Forensics (Orient: the Cyber Response Team [CERT][12] or other organization recovers the data and gets the full picture, and if necessary, the intelligence community analyses and evaluates it), Judgement in organizational process (Decision-1: The competent authorities, based on their assessment, draw up a plan for dealing with the cyberattack in question, taking into account the legal interpretation of the attack and the political needs of both the international and domestic arenas), Decision-making (Decision-2: a political decision on how to deal with the attack, from ignoring it to retaliating against it), and then Retaliation (Act), if necessary. The initial problem is that this process involves numerous entities, including the private sector, and there are multiple branches, so it will require a considerable amount of time to complete.

Moreover, an additional concern pertains to the intricacies inherent within each process. The second problem is to the realm of digital forensics. The forensic process is characterized by three distinct difficulties. Two of these are of a technical nature. First, the sheer volume of data has increased exponentially. Actually, the capacity of hard drive storage at nearly all PCs has expanded from megabytes to terabytes. Consequently, even the creation of a forensic disk image and the performance of an investigation require a significant amount of time. Second, the qualitative difficulty of data retrieval and analysis has increased with advancements in virtualization and encryption technology, the Internet of Things (IoT) and code obfuscation, resulting in the distribution of data and obstruction of investigations.

Third, attacks by state or non-state actors have become more sophisticated. In 2011, Eric M. Hutchins of Lockheed Martin and his colleagues proposed "a systematic process to target and engage an adversary to create desired effects,"[13] known as "the cyber kill chain." It comprises seven steps (see Table 1). Attackers use every possible tactic to avoid detection, or if detected, to preserve anonymity. Concretely, they employ (a) route camouflage (e.g. using TOR[14] or multi-hop VPNs), (b) attacker disguise (spoofing IP addresses, or impersonating dummy companies, hackers, or soldiers), (c) false origins (using "botnets,"[15] made up of hacked servers), and (d) concealing intent. In the 2022 Russian invasion of Ukraine, a Hacker group, so-called "IT ARMY" was widely reported—arguably an example of an attacking actor mobilizing volunteer hackers by ideological incitement, further complicating identification of the real attacker.

The third problem pertains to the decision-making process within the government (Decision-1). In this process, each ministry and agency must interpret both domestic and customary international laws. Moreover, they must establish a set of criteria for determining whether situation constitute an armed attack, and retaliation is appropriate. However, in the absence of a well-developed process and legal interpretation, they may encounter difficulties in formulating a proposed response policy that aligns with the legal interpretation of cyberattacks. Ultimately, the government must reach a political decision on a 10-step scale from "ignore" to "retaliation,"[16] as articulated by Jason Healey. This decision must be made with consideration for both domestic and international political factors. However, the absence of clearly defined standards and the ambiguity of responsibility leave no option but to adopt a "doing nothing." In case the retaliation is an option, the question of proportionality in relation to the damage caused by the attack back and the compensation for damage of side effects to third parties will also need to be considered. It is evident from the counterattack cases described below that in the U.S. and the U.K. there are instances where the true attacker is not disclosed against an overseas

source. It can be estimated that this process is to a considerable level already established in both countries. Nevertheless, they recognize the need to further strengthen the system and expedite responses.

The fourth challenge relates to the efficiency and clarity of the decision-making process in the political authorities (Decision-2). In 2017, James E. McGhee outlined challenges pertaining to U.S. OCO decision-making, compared with that for physical attacks.[17] First, launching OCO requires the approval of the Secretary of Defense or the President—senior authorities—and thus may lose vital speed. Second, if the counter OCO is still preparing the early phases of the cyber kill chain (Table 1), it can be unrealistically time-consuming from phases 1 through 6 compared to physical attacks. This underscores the need for delegated authority and for operational planning that synchronizes OCO with kinetic actions. Otherwise, the required speed and the operational effectiveness may be compromised.

The fifth challenge is that of retaliation (Act), especially cyber counterattacks. As Timothy M. McKenzie contends that "cyberattacks are inherently vulnerable"[18] because once a counterattack is launched, the infrastructure used for that attack (e.g. botnets, command-and-control systems, etc.) and any exploited vulnerabilities are exposed, rendering the other ones for any future attacks. Furthermore, the continuous defense's Cyber Intelligence, Surveillance, and Reconnaissance (hereinafter [CISR]), such as the continuous intelligence gathering on vulnerabilities, hidden infiltration routes, and vital data is needed. The credible and continuous "preposition" for OCO, that is, an infrastructure for cyber counterattacks, which is preparation and maintenance for the counterattack, or infiltration infrastructures by armed forces or intelligence organizations necessitates for ensuring cyber deterrence capability because it faces a persistent risk of detection and remedial action by the adversary's CISR.

(3) Deterrence Stemming from Inter-State Interaction and Associated Challenges

Some scholars note that international relationship between attacking country and defending country can themselves yield a certain measure of deterrence. In 2017, Nye posited that although traditional deterrence might still play a significant role in cyberspace, its effect is inherently limited, so he proposed two additional types of deterrence. One is deterrence by entanglement in which states with deep interdependency do a cost-benefit calculation and choose not to attack each other. The other is deterrence by norms and taboos, from which assumes an attacking state with a high regard for its reputation will be deterred by the risk assessment of the reputational damage it would suffer if exposed. Nye also sees potential for moderate improvements in deterrence by punishment through the ongoing refinement of ICT and permitting attribution with certain ambiguity, which enable government to react with confidence. He is convinced that the combination of these new forms with traditional deterrence can help mitigate cyber threats.

In 2021, Stephan Soesanto definitively classified six types of deterrence:[19] (1) deterrence by denial, (2) deterrence by punishment, (3) deterrence by entanglement is the same as Nye's definition, (4) deterrence by delegitimization is based on the idea of deterring attacks by imposing reputational damage on the attacker through statements of condemnation of violations of norms as same as Nye's "norms and taboos." This approach

encompasses the idea of public attribution (PA), which Kristen E. Eichensehr defines as the act of publicly identifying the malicious tools, the perpetrators, and the responsible adversary behind an attack.[20] (5) Deterrence by reputation is a clear and effective method of discouraging adversaries by demonstrating credible offensive cyber capabilities and making them evaluate them. A prominent example of this is the Stuxnet attack (attributed to the U.S. and Israel) on Iran's nuclear facilities in 2009. (6) cross-domain deterrence is an approach that the U.S. has claimed may be used in response to cyberattacks, including kinetic strikes.

It is imperative to emphasize that both Nye and Soesanto have underscored the fact that each form of deterrence can be effective under particular conditions. They have also expressed the view that traditional deterrence still plays a certain role. However, it differs from this research in that it does not appear to have closely examined ACD, HFO, or OCO as deterrence power in gray zone environments, which have evolved significantly in recent years.

(4) Directions for Improving Cyber Deterrence

There are three broad approaches that may be considered to address the attribution problem. First, the challenges of digital forensics should be considered. Whilst it is acknowledged that further advancements in forensic technology are necessary, it is also important to establish a certain threshold[21] of confidence that will enable policymakers to make a decision. Furthermore, public attribution, whereby a high level of certainty identifies a specific state or state-linked actor, can enable states to impose retaliatory measures, such as official condemnation or economic sanctions, thereby contributing to deterrence by punishment.

In recent high-profile cases of malicious software,[22] such as WannaCry and NotPetya, a so-called attribution coalition,[23] was initiated by the Five Eyes (the U.S., the U.K., Canada, Australia, and New Zealand). This coalition publicly released information naming North Korea and Russia (and actors under their influence) as responsible for indiscriminate attacks on public and private targets and collectively condemned them. This joint disclosure is likely to have had a certain deterrent effect on potential future attackers and to have served as a warning to the states behind these activities.

Second, regarding the absence of a coherent framework for national attribution processes: in the U.S., a triad of the Department of Defense, the Department of Justice, and the Department of Homeland Security—plus the Defense Industrial Base and the National Cybersecurity and Communications Integration Center—underpins smooth and interactive information sharing and response when critical infrastructure is targeted. In the U.K., Government Communications Headquarter (GCHQ) acts as a control tower, with the NCSC operating under its oversight. The NCSC shares intelligence with other agencies and assists with information sharing for the private sector. This integrated arrangement is intended to ensure swift responses to cyber incidents and represents an efficient framework.

The third issue pertains to the execution of counterattacks. As previously mentioned, phases 1 to 6 of the cyber kill chain (see Table 1), are imperative for OCO, including the

platform maintenance that ensures infiltration routes remain accessible for subsequent operations. This necessitates a more profound integration of military and intelligence operations, accelerating the OODA loop and, by extension, reinforcing the role of OCO as a mechanism for deterrence through punishment.

2. The Transition of U.S./U.K. Deterrence Posture (with HFO, ACD, and OCO) and Case Studies

It is estimated that China's cyber strategy is aimed at achieving primacy in information technologies, including data theft, manipulation and destructive acts, which are intended to serve its national interests. Furthermore, Russia is believed to be pursuing an aggressive informatization agenda, which is believed to be accompanied by psychological operations. In response, the U.S. and the U.K. have adopted a focus on defensive measures for the purposes of either deterrence or prevention. This chapter conducts an examination of the evolution of the U.S. and the U.K. cyber strategies, identifies the direction of their approaches to cyber deterrence, and examines them with case studies of ACD, HFO, and OCO to verify the resulting effectiveness in cyber deterrence.

(1) The U.S.

In the 2022 National Defense Strategy, the U.S. introduced the concept of integrated deterrence, which encompasses crises and gray zone contingencies by employing every tool at its disposal, including diplomatic, informational, military, and economic means. These tools are employed in collaboration with allies and partners across all five operational domains. Since 2010, the U.S. has established the USCYBERCOM, and it has been charged with the responsibility of implementing self-defense rights under the customary international law in the cyber domain. Furthermore, the command has been authorized to conduct OCO, as defined as projecting power[24] in and through foreign cyberspace through actions taken in support of CCDR (Combatant Commander) or national objectives.

In 2012, the U.S. Department of State Legal Adviser Harold H. Koh said during an interagency conference hosted by the Department of Defense, “Cyber activities may in certain circumstances constitute uses of force within the meaning of Article 2(4) of the UN Charter and customary international law” and reaffirmed that “If the physical consequences of a cyberattack work the kind of physical damage that dropping a bomb or firing missile would, that cyberattack should equally be considered a use of force.”[25]

In essence, the policy of the U.S. positioned cyber counterattacks, akin to self-defense, within the ambit of customary international law. Meanwhile, the 2011 Department of Defense Strategy for Operating in Cyberspace defined ACD as a synchronized real-time capability to detect, analyze, and address threats and vulnerabilities in order to limit damage. This definition also pertains to dealing with below gray zone situations.[26] At that time, the concept of cyber deterrence was predicated on a denial-centric approach[27] to the issue. However, as Martin Libicki contended, the use of deterrence by denial in isolation is significantly disadvantageous when evaluated from a cost-benefit perspective.[28]

It is evident that the 2018 Department of Defense Cyber Strategy underscored the significance of safeguarding Department of Defense networks (DoDIN), U.S. critical national infrastructure, and allied and partner nations from cyberattacks that reach the level of using force. To this end, the strategy emphasized that the Department of Defense will take measures to deter hostile cyber operations and maintain capabilities to respond with any necessary means should deterrence fail. In the context of this initiative, HFO have been initiated. The implementation of these operations entails the U.S., in conjunction with allied or partner nations, installing detection equipment on the systems of those nations' territory.

The primary objective of this initiative is to identify potential cyber threats at an early stage, with the aim of preventing or disrupting them before they reach the U.S. The U.S. regards ACD and HFO both as part of Defensive Cyber Operations (hereinafter [DCO]), though HFO includes aspects such as disrupting or blocking at the source,[29] making it partly akin to OCO. From 2018 to March 2023, the USCYBERCOM executed 40 instances of HFO across 21 nations, encompassing 59 networks.[30] It is asserted that the objective of these operations is to enhance the resilience of the targeted nation's networks, as well as that of the U.S. homeland, through the dissemination of threat intelligence. The Department of Defense's 2023 Cyber Strategy document reaffirm the continuous development of HFO.

a. Verification of HFO

(a) From December 2021 to March 2022, the Cyber National Mission Force (CNMF) of USCYBERCOM cooperated closely with the Ukrainian government to carry out its most significant HFO mission. This initiative entailed the CNMF collaborating with Ukrainian cyber personnel to identify, address, and neutralize or mitigate potential threats to critical government networks.

The Public Affairs Section of the CNMF has announced[31] that the organization has proactively identified potential threats and provided vital intelligence to Ukraine in order to strengthen the latter's resilience. In addition, the CNMF has distributed this intelligence to the U.S. in order to bolster defense before threats reached homeland systems. The question of whether there was any “disruption at source” activity in this instance remains unclear.

(b) In May 2022, an HFO in Lithuania concluded that, following monitoring of cyber activities in the networks of the Lithuanian Ministry of Defence and Ministry of Foreign Affairs by US-CNMF personnel in collaboration with local authorities, there was evidence of unauthorized cyber activity. Margiris Abukevičius, Lithuania's Deputy Minister of Defence, emphasized that the HFO has ensured that, regardless of whether the country is in a state of peacetime or wartime, critical infrastructure must be protected in advance. In a similar context, the CNMF website states, not just for hosted nation but for the U.S. homeland defense, we are able to obtain adversary the Tactics, Techniques and Procedures (TTPs) before an attack is launched.[32]

In order to obtain TTPs of the adversary prior to an attack, it is possible to strengthen the resilience of the DCO by updating virus definitions and so forth. Furthermore, this process

will provide essential information in terms of conducting OCO to obtain the adversary's vulnerabilities and their tactics. Consequently, it contributes DCO, and CISR for OCO activity. However, the question of whether this constitutes actual disruption at source remains unanswered in these public statements.

b. Verification of OCO

(a) In November 2021, CNN reported that the commander of USCYBERCOM, General Paul M. Nakasone, admitted that the U.S. had conducted counterattacks[33] "retaliatory strikes" over the past three months to punish the hackers who launched a ransomware attacks on Colonial Pipeline. The targets were believed to be the facilities of entities that had provided funding to the ransomware groups, many of which were in Russia and Eastern Europe.

This case is of particular significance as it represented official acknowledgement that the U.S. carried out OCO abroad in a situation that did not constitute armed conflict—i.e. in gray zone—against a hacker collective. With the exception of the systems and methods employed by hacker groups to infiltrate target networks and the subsequent actions of the U.S. in neutralizing them, specific details remain undisclosed. Nevertheless, the U.S. has demonstrated a willingness to retaliate whenever a cyberattack crosses a certain threshold. This behavior serves as a signal to potential opponents that they risk significant retaliation if they act in a confrontational manner. Consequently, while deterrence by denial dominated in the early 2010s, the U.S. has started to place greater emphasis on deterrence by punishment, to use OCO as a deterrence power in gray zone situations.

(b) In June 2022, CNN reported[34] that General Nakasone revealed that the U.S. had conducted OCO to assist Ukraine in defending itself from Russia's invasion. He stated that a variety of operations had been conducted, including both offensive and defensive maneuvers, as well as information operations. Despite the absence of explicit details, the Biden administration, cognizant of the potential for escalatory retaliation by Russia, had committed to refraining from military engagement. Notwithstanding, this constitutes the first instance in which the U.S. has acknowledged and executed OCO against a nation not to transgressing beyond the bounds of armed conflict, as delineated within gray zone.

In response to a query regarding the potential for inconsistency White House press secretary Karine Jean-Pierre said, "We don't see it as such," when asked at a news conference whether Cyber Command's actions contradicted Biden's pledge. It could be argued that cyberattacks do not constitute military engagement directly. The administration deems them permissible up to a certain intensity. The same media source also cited a senior official at the Defense Intelligence Agency hypothesizing that Russia "has not carried out a clear retaliatory cyberattack" because it seeks to "avoid further escalation and focus on its improvement of national defense capability." Therefore, from the perspective of deterrence by punishment, and also in accordance with Soesanto's concept of deterrence by reputation, it can be assumed that the U.S. is willing to conduct OCO in gray zone circumstances, thereby exerting a powerful deterrent effect on nation-state adversaries.

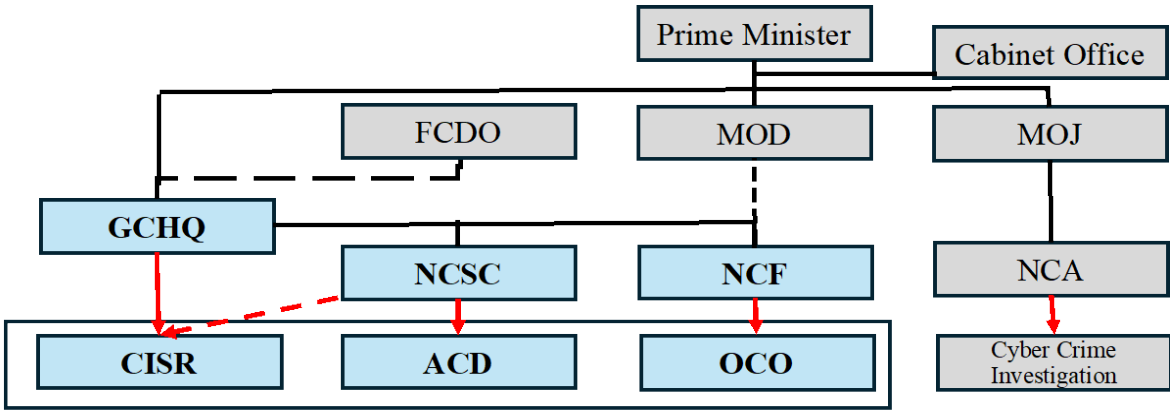
(2) The U.K.

The U.K. is the world's third largest cyber power.[35] As early as June 2009, the then-Minister for Security in the Home Office, Lord Alan W. J. West, acknowledged that GCHQ possessed the capability to counterattack hostile computer systems,[36] albeit unofficially and as a unstated reference. Subsequently, in 2011, the U.K. published its National Cyber Security Strategy, designating the Cabinet Office, GCHQ, the Ministry of Defence, and the newly established National Crime Agency (NCA; under the Home Office from 2013) as the country's cyber nucleus, consolidating dispersed authorities. It was anticipated that GCHQ would assume a leading role in the field of operational cybersecurity.

The 2016 National Cyber Security Strategy represented a significant expansion of the previous strategy, with funding increasing from £850 million to £1.9 billion. It was anticipated that over a dozen cybersecurity organizations would be integrated into the NCSC under GCHQ, with the objective of centralizing intelligence collection and incident response. Furthermore, the strategy explicitly states the implementation of OCO, whereby the U.K. would detect, understand, investigate, disrupt, and pursue attackers, and, where necessary, undertake offensive action.[37] The 2022 National Cyber Strategy adopts a comprehensive and integrated approach, emphasizing "sustained and tailored deterrence" and providing flexible, adaptable responses according to threat level. In the aforementioned document, ACD is characterized by the proactive identification and remediation of vulnerabilities, the response to incidents, and the automatic interruption of attacks in and through cyberspace. In pursuit of this strategy, it is imperative that the full spectrum of capabilities is employed.

The greatest changes have been the expansion of the NCSC's capabilities and the establishment of the NCF in 2020 (see Figure 2). The NCSC serves as the national information hub for cybersecurity, handling major incidents and providing a safer internet environment to citizens and enterprises. It coordinates with the newly formed NCF, which conducts more direct countermeasures.

Additionally, the NCSC is strengthening the role of GCHQ, particularly in collecting intelligence and attributing attacks, in partnership with other intelligence agencies and Five Eyes countries, thus improving the detection and identification of malicious cyber actors. Under the same national cyber strategy, the NCSC is described as having world-class capabilities in cyber-threat detection, assessment, analysis, and ACD, functioning as one of the intelligence agencies, delivering crucial capabilities to enable national-level attribution of major incidents, which underpins cyber deterrence. In essence, policymakers can make their decisions to conduct OCO operations based on the principles of proportionality and necessity, even in circumstances where precise attribution is arduous and time-consuming. This approach serves to mitigate the attribution problem. Finally, with regard to the conduct of OCO operations abroad, NCS 2022 has included the stipulation that the NCF must adhere to strict scrutiny under the Intelligence Services Act 1994 and the Investigatory Powers Act 2016 to maintain oversight and protect communications secrecy, thus preventing a repeat of the Snowden affair.



Examples: 1. Solid lines are indicated to demonstrate command and control relationships, whilst dotted lines are used to denote legal command relationships. Solid arrows are employed to signify primary responsibilities, and broken arrows are used to indicate partial implementation responsibilities.
FCDO: Foreign, Commonwealth and Development Office

Figure 2: The UK's Cyber Framework

(Created by the author based on UK Ministry of Defence materials.)

a. Verification of ACD

(a) In an interview with the Nikkei in July 2024,[38] Chief Executive Officer of the NCSC, Felicity Oswald, commented that "through ACD, we are proactively discovering threat intelligence and similar data, removing over two million malicious websites, and collecting intelligence in a proportionate and transparent way." She noted that the Investigatory Powers Act 2016 clarifies what intelligence agencies can do to ensure that any interception of private communications is always subject to oversight, and then that protecting the security of the nation is paramount. This suggests that the NCSC, as an intelligence agency, can lawfully collect information - including from communications not explicitly targeted - take down malicious websites, and otherwise intervene where justified.

(b) The website of the NCSC[39] states that the organization provides various security tools and services to domestic businesses upon request, via its "MyNCSC platform." These tools offer functionalities such as early warning systems, mail checks, and web checks. Government departments can similarly adopt the NCSC's cyberattack disruption solutions by installing them on their systems. Consequently, the NCSC, operating as an integrated hub, concurrently undertakes data collection to eliminate malicious websites, while distributing the requisite information and services to both public and private entities. As previously mentioned in point (a), ACD has been demonstrated to enhance resilience, which is a fundamental component of deterrence by denial.

b. Verification of OCO

In April 2018, Jeremy Fleming, who was Director of GCHQ at the time, stated at the CYBERUK conference[40] organized by the NCSC that, with the cooperation of the Ministry of Defence, the U.K. had conducted cyberattacks on the so-called Islamic State of Iraq and Syria, causing "disruptive damage" to its online presence. The U.S., Australia, and France also admitted to equivalent operations. Thus, it is certain that the U.K. has

undertaken OCO against a non-state adversary in gray zone below the threshold of armed conflict.

Whether this was carried out by the NCF is not specified, but such activity would be legal under the Investigatory Powers Act, suggesting that in cost calculations by non-state attackers, the risk of the U.K. retaliation is considerably elevated. This suggests a substantial deterrent effect in the deterrence by punishment realm.

3. How the Development of ACD, HFO, and OCO in Gray Zone Affects Traditional Deterrence — These Impacts and Limitations

As outlined the U.S. and the U.K. strategies and the operational cases in the last chapter, it is evident that from the latter half of the 2010s onward, measures such as ACD, HFO, and OCO align with the directions of integrated deterrence and comprehensive strategy and have tangibly enhanced traditional deterrence.

Two broad features emerge from this analysis. First, OCO in gray zone may have significantly altered the landscape for deterrence by punishment. The term OCO was formerly regarded as a military operation for wartime or serious hostilities. Therefore, it was not disclosed for below armed attack situation, such as gray zone situations. Consequently, in gray zone operations where attacks are frequent, the attacker's awareness that if we launch a cyberattack, we may actually face an OCO response, which had not been firmly established. Further discussion is required on the interpretation of customary international law, announcing the employment of OCO as a deterrence power in gray zone situations obliges attackers to consider the risk of genuine retaliation. That risk raises their apprehension of a counterattack in risk-benefit evaluation and thus serves as a potent cyber deterrence by punishment.

Second, the fact that OCO can be deployed in gray zone circumstances implies that the U.S. or the U.K. is continuously engaged in CISR, i.e. gathering intelligence on vulnerabilities and infiltration routes to maintain preposition for OCO capability (phases 1–6 of the cyber kill chain, Table 1). Again, this compels potential attackers to consider in the high probability of retaliation. In other words, an adversary is forced to recognize that if he initiates an attack, he may face immediate countermeasures affecting deterrence power overall.

However, an issue remains unresolved: Is OCO in gray zone lawful under customary international law? As Masahiro Kurosaki notes, Japan’s ACD, what this research calls OCO in gray zone, is subject to the prohibition of the use of force, the principle of non-intervention, and the principle of State sovereignty in customary international law. Whether it violates these principles depends heavily on the specific circumstances, and states have different interpretations.[41] Referring to the Tallinn Manual, Kurosaki points out, “There is no established international consensus on what constitutes the ‘use of force’ in contexts that do not involve direct physical destruction or casualties.”[42] To implement this overseas, therefore, it is essential that national legislation is enacted to give intelligence agencies and the Japanese Ministry of Defense (JMOD) the authority to use capabilities in times from armed conflict to peacetime, and to clarify the interpretation of customary international law. Accordingly, the U.S. interprets the right of self-defense

under customary international law to authorize OCO, regarding the authority under Title 10 and Title 50 of the U.S. Code, allowing the NSA and USCYBERCOM to act in accordance with already established Rules of Engagement. And the U.K. grants GCHQ and the NCF authority to conduct OCO. However, as there is no consensus in the international community on the thresholds for OCO, further discussion is needed.

4. Comparison of Japan, the U.S. and the U.K.

(1) The Process from the National Security Strategy to the Enforcement of the Japan's ACD Acts

In 2022, the National Security Strategy[43] explicitly stated for the first time that Japan would possess ACD capabilities. As a crucial step toward realizing this, on May 23, 2025, the Acts on the Strengthening of "Cyber Response Capabilities and Related Measures (hereinafter Japan's ACD)" were promulgated,[44] thereby establishing the legal framework for ACD—a long-standing issue. Needless to say, this signifies a major shift away from a purely defensive posture—that is, from a deterrence strategy reliant solely on the resilience of systems—and, with enforcement scheduled for October 2026, preparations to build these capabilities are currently proceeding with great urgency.

There are four major areas where significant improvements are expected as a result of these legislative reforms: organizational and structural reforms, the implementation of neutralization measures, the strengthening of public-private partnerships, and the promotion of the utilization of communications data.

At first, organizational and structural reforms; this addresses the long-standing cybersecurity issue of the absence of a central Headquarters. The National Cybersecurity Office (hereinafter "NCO"), headed by the National Cyber Director—a position held concurrently by the Deputy Director of the National Security Secretariat (NSS), which succeeded the National Information Security Center (NISC)—has been established across both the Cabinet Secretariat and the Cabinet Office. Under the "strong comprehensive coordination" provided by this NCO, the JMOD and the National Police Agency (NPA)—the ministries responsible for implementing measures—have become legally authorized to carry out neutralization measures. It can be understood that, when deemed necessary from the perspectives of irreplaceability and urgency, and while taking into account proportionality, communication protection measures[45]—referred to in this paper as OCO—can now be executed.

Second, in terms of neutralization measures, following the enforcement of the Japan's ACD Acts (one year and six months after its promulgation), after obtaining approval from the Cyber Communications Information Oversight Commission—which ensures the proper use of communications information[46]—the "Cyber Threat Prevention Measures Executive Officer" of the NPA will implement neutralization measures pursuant to Article 6-2 of the amended Police Duties Execution Act. However, according to the government's view, 99% of harmful communications originate from overseas (so-called "overseas-to-domestic" and "overseas-to-overseas" communications). While urgent situations posing a risk of serious harm may involve potential criminal offenses under the Penal Code, it is presumed that the Ministry of Defense and the Self-Defense Force are highly likely to take communication protection measures by applying the Police Duty Execution Act

mutatis mutandis to communications related to national security and diplomacy. In such cases, the operational units, which should be the Japan Self-Defense Force Cyber Defense Command, will implement these measures in cooperation with the police; however, following consultation with the Ministry of Foreign Affairs,[47] they will apply the Police Duties Execution Act mutatis mutandis and carry out the neutralization measures through procedures similar to those followed by NPA enforcement officers. However, ex post facto reporting is permitted in cases where there is no time to act otherwise.

Regarding the third point—strengthening public-private cooperation—the establishment of the "Council on Information Sharing and Countermeasures" was mandated, and a reporting obligation was explicitly stipulated for Critical National Infrastructure (CNI) entities[48] within the cooperative framework designed to facilitate attack response. Consequently, reporting to the government by CNI entities and others in the event of a cyber incident, which was previously voluntary, has become mandatory, enabling the government to obtain information more quickly.

Lastly, the use of communications data; the secrecy of communications guaranteed by Article 21 of the Constitution had been a major issue. However, government agencies are now able to obtain and use information deemed necessary for crime prevention and national security after undergoing an audit of appropriateness, making it legally possible to intercept communications related to malicious cyber activities either ex post facto or in advance.

As a result of these legal reforms, how effectively can we address various situations? The scenarios where actual measures are required can be guided by Kurosaki's observations. Specifically: (a) Implementation is warranted in cases where there is a risk of a serious cyberattack that raises security concerns for Japan, though not amounting to an armed attack; (b) identifying (through information gathering and analysis) the perpetrators of the potential attack, the servers and other means of attack to be used, as well as their locations; (c) eliminating or neutralizing the target using cyber means (information and communication technology); and (d) intruding (presumably via remote control) into computers, the Internet, and other networks or systems as necessary.[49]

In this Kurosaki scenario, specifically, when comparing the capabilities of cyber-advanced nations such as the U.S. and the U.K.—which combine OCO and ACD capabilities—with the capabilities anticipated under Japan's ACD Acts, it becomes clear that Japan still faces several challenges to conducting neutralization. These include issues related to organizational structure, operational effectiveness, and intelligence organizations that fall outside the scope of this law—all of which are expected to act as bottlenecks in actual implementation.

To clarify these challenges, it will be compared that Japan's approach with the initiatives of the U.S. and the U.K., which served as significant references in drafting the Japan's ACD acts.

In other words, to implement neutralization measures—which are an extension of cyber attack response—it will be conducted that a comparative analysis in the following

sections based on six items that comprehensively cover the required functions or capabilities. That is, (a) The effectiveness of command and control within the government organization deploying ACD; (b) The speed of the OODA cycle, including public-private partnerships; (c) CISR (mainly, cyber espionage) capabilities to identify the source (or perpetrator); (d) Preparation and maintenance of the infrastructure for cyber counterattacks (preposition) and deployment capabilities; (e) a government framework for aggregating not only cyber intelligence but also more comprehensive intelligence; and (f) OCO case studies in coordination with PA to enhance the credibility of retaliation (counterattack).

(2) Current Status the U.S.

a. While OCO is led by USCYBERCOM and intelligence collection is primarily handled by the NSA; both organizations are under the command of a single individual (the Commander of USCYBERCOM and Director of the NSA).

b. Regarding public-private cooperation, not only have two-way information-sharing frameworks such as the Defense Industrial Base, but the Director of the NSA reports directly to the Secretary of Defense and the President.

c. In cases where significant harm related to cyber incidents to the lives and property of the public is anticipated, it is presumed that the NSA will primarily conduct intelligence collection (more specifically, cyber espionage) activities to the extent necessary (within the scope of exercising the right to self-defense).

d. Although the actual details have not been disclosed, as can be inferred from the examples mentioned above, the U.S. possesses the capability for the NSA or USCYBERCOM to create and maintain the preposition necessary to carry out OCO.

e. Intelligence agencies such as the NSA, Central Intelligence Agency (CIA), Defense Intelligence Agency (DIA) and other 14 organizations possess the capability to provide intelligence that supports comprehensive decision-making, including information beyond the cyber domain.

f. Since the 2017 WannaCry incident, the U.S. has disclosed information either independently or in collaboration with the Five Eyes countries, aiming not only to achieve certain effects in terms of deterrence, improved defensive capabilities, and norm-setting,[50] but also to make both state and non-state actors assess the risks of retaliation by carrying out counterattacks in collaboration with such disclosures against them. In other words, it makes both actors consider the credibility of retaliation much higher.

(3) Current Status in the U.K.

a. GCHQ takes operational command over the NCF, and together with the intelligence divisions within GCHQ, a unified command structure is in place.

b. The Director of GCHQ reports directly to the Prime Minister. Regarding public-private cooperation, a two-way information-sharing system has been established with the NCSC to the private sectors openly, under GCHQ, serving as the hub information center.

c. While complying with the Investigatory Powers Act and the Intelligence Services Act, it possesses the capability to collect necessary intelligence within the scope of conducting OCO operations.

d. Although the actual details have not been disclosed, based on its capability to carry out counterattacks against ISIS, it is presumed to possess the ability to launch counterattacks using preposition held by such organizations.

e. GCHQ is in a position to oversee other intelligence agencies, enabling centralized and comprehensive intelligence aggregation.

f. As a member of the Five Eyes alliance, it conducts PA against both state and non-state actors and has a track record of conducting OCO against non-state actors, possessing capabilities and intentions similar to those of the U.S.

(4) Comparison with Japan After the Law's Enforcement and Related Challenges

Below, challenges will be clarified by comparing the organizational structure and capabilities anticipated in Japan after the Japan's ACD Acts enforcement with those of the U.S. and the U.K.

a. The Merits and Demerits of Centralized Command and Control within the Government

The NCO, operating under the Cabinet Secretariat and Cabinet Office, is expected to serve as the command center for ACD in consultation with the NSS; however, it is mandated to do "coordinate strongly and comprehensively" with relevant ministries and agencies.

As is evident from the organizational structures of the U.S. and the U.K., quick response and decision-making capabilities, and precise target identification to avoid accidental attacks on third parties are essential for effective neutralization measures. To achieve this, a centralized command and control system is necessary—one in which intelligence agencies and counterattack deployment agencies are integrated, such as a command headquarters like the U.K.'s GCHQ that operates the NCF, the operational entity of the OCO, in a unified manner. Furthermore, since the Cyber Defense Command, which serve as the operational body for neutralization, is under the command of the Minister of Defense, and the National Cyber Unit is under the command of the Commissioner General of the NPA, complex coordination would be the barrier of Japan to quickly determine whether a major cyber incident constitutes a national security issue or a cyber-crime issue, and to decide the extent of the neutralization measures to be taken. According to the "Explanation of the Japan's ACD Acts," by Cabinet Office, the two ministry and agency are required to "cooperate closely, such as by working in the same building" when carrying out their missions. However, the relationship between the NCO (or NSS) and these organizations cannot be interpreted—either legally or from the phrasing "strongly comprehensive coordination"—as allowing for substantive command, and it is unclear whether neutralization as a counterattack can be carried out swiftly.

Adding to those, to carry out missions in a unified manner, as same as in the U.S. and the U.K., it is desirable to operate under a single command. However, it is unclear neither there are a single individual or two individuals to whom command is delegated at the operational level, nor are their respective responsibilities defined. It is desirable to establish a system where personnel carry out missions together and can determine whether the situation pertains to national security or criminal law based on the nature of the incident.

As a summary of the organizational structure, Figure 3 shows a conceptual diagram of the information flow and chain of command in the event of a cyberattack, based on a diagram from the Cabinet Office.

b. Functionality of the OODA Cycle, Including Public-Private Partnerships

After a CNI entity reports a major incident to the NCO and the relevant government agency without delay, the report is forwarded via the NCO to the NPA and the JMOD, which are responsible for deploying the ACD. Following forensic analysis and other procedures, a report on the affected server—based on Indicators of Compromise (IoC) and TTPs—is submitted to the Prime Minister via the NCO. Subsequently, instructions regarding the deployment of the mission are issued to the NPA and the JMOD via the NCO.

However, it is not explicitly stated that this involves a two-way information-sharing mechanism or system similar to the U.S. Defense Industrial Base. Sharing information necessary for the system defense of CNI entities enhances deterrence; therefore, it is necessary to diligently establish a framework that contributes to system resilience, such as the early remediation of discovered vulnerabilities. Furthermore, a framework for two-way information sharing between the public and private sectors does not currently exist, and discussions are scheduled to begin in the fall of 2026 or later.

c. Intelligence Capabilities for Identifying Perpetrators
It is difficult to identify a true perpetrator behind an attack based solely on expert analysis from a technical perspective, such as IoC and TTPs extracted through digital forensics. To identify true attackers who are the targets of mitigation measures, cyber espionage—which involves tracing the source and gathering necessary information—is indispensable; however, it remains unclear which government agencies have been specifically authorized to employ such fundamental intelligence gathering, that is, almost as same as cyber espionage methods.

In the U.S. and the U.K., intelligence agencies such as the NSA and the GCHQ conduct cyber espionage, operating under a system integrated with CYBERCOM and NCF, which are the respective operational force of OCO.

In Japan, there had previously been significant restrictions—such as obtaining information only from the server immediately adjacent to the compromised server—until the Cybercrime Act made it legally possible to trace the source of a cyber incident. To ensure a swift, appropriate, and accurate response, it is necessary to designate or establish an agency with centralized authority.

d. Preposition Capabilities for Cyber Counterattacks
Since the possession of counterattack capabilities was not originally anticipated, the state could not establish a legal basis for conducting such activities; consequently, Japan does not currently possess such capabilities. Building these capabilities is necessary moving forward.

e. A Government Framework for the Comprehensive Aggregation of Information Beyond Cyber Intelligence
Regarding decisions on neutralization measures, while it is necessary to determine measures that take proportionality into account after satisfying the requirements of urgency and non-substitutability, it is often difficult to identify the true attacking entity behind an attack based solely on expert opinions from a technical perspective, such as IoC and TTPs extracted from digital forensics. Furthermore, while cyber espionage is an important means of uncovering attack networks (botnets) and identifying true attackers, there is a possibility that the identified entity is not a genuine perpetrator or that another state actor is operating behind the scenes. To identify these actors, it is necessary to comprehensively analyze all available information—including OSINT (open source intelligence), satellite imagery, HUMINT (human intelligence), and profiles of individuals involved in the attack etc.—to raise the probability of identification to a level where the government can make a decision.

To propose a government agency responsible for decision-making in Japan, the National Cyber Director of NCO—who also serves as Deputy Director of the NSS—would be the entity to provide substantial insights into international politics. Furthermore, the NPA or JMOD must conduct forensic analysis, and other intelligence agencies must provide all information other than cyber. Ultimately, the NCO will be responsible for obtaining the Prime Minister's decision.

However, while the NCO possesses the ability to coordinate information gathering for government decision-making—that is, political decisions—it is not an organization that has the authority which possess authority over the aggregation of information and comprehensive decision-making.

In the U.S., regarding major cyber incidents, the NSA and the USCYBERCOM operate under the one commander / director, enabling information sharing between the two organizations. More importantly, while information sharing within the intelligence community—including the NSA, CIA and DIA—occurs as needed, the three major cyber-related organizations—the DoD, Department of Homeland Security, and Department of Justice—divide their roles and share the necessary information regarding ACD or OCO to ensure swift and accurate decision-making.

Furthermore, in the U.K., since GCHQ serves as both the information aggregation hub and the entity responsible for deploying countermeasures, when necessary, it is capable of providing comprehensive intelligence for political decision-making and launching rapid counterattacks as an authority.

Japanese Prime Minister Sanae Takaichi has proposed the establishment of a National Intelligence Agency (still pending as of May 2026) within the Cabinet Office. If a system for consolidating intelligence is adopted, and an organization is created that brings together the authority and personnel capable of comprehensively evaluating both cyber intelligence and other forms of intelligence, it would be possible to provide proposals that significantly contribute to the deployment of ACD.

It is likely necessary to establish an independent intelligence-gathering framework capable of objectively assessing the validity of information—not relying solely on intelligence from the U.S. and other friendly nations.

f. OCO Performance Linked to PA Capabilities to Enhance the Credibility of Deterrence

The NPA already implemented this in 2017 following the large-scale WannaCry attack. It is desirable that not only the NPA but also the JMOD proactively conduct PA regarding cyberattacks related to national security. Furthermore, if we intend to pursue closer cooperation with the Five Eyes nations in the future, it is desirable to establish an information-sharing framework for PA with the NCO serving as the coordinator (or commanding authority).

The results of the comparison of the above items are summarized in the Table 2.

Therefore, the issues requiring improvement in the future are as follows.

a. A government organizational structure centered on an organization that consolidates the information necessary for ACD implementation and possesses the authority to carry it out

b. A two-way organizational structure encompassing both the public and private sectors that enables rapid OODA loops and securing their systems

c. A integrated organizational structure for cyber intelligence activities

d. The need for legal frameworks to grant authority to the NCO (or JMOD) to operationally command the Defense Intelligence Headquarters or Cyber Defense Command in order to execute preposition activities

e. The need for an integrated intelligence framework that aggregates information comprehensively under single authority in the intelligence group, not limited to cyber

f. Advancement of international cooperation for PA aimed at greater effectiveness

To summarize, in order to effectively implement ACD, Japan must establish a system centered on an organization with single authority to centrally aggregate both cyber-origin intelligence—which the nation has historically lacked—and all other intelligence-gathering capabilities. This does not necessarily mean merging all organizational functions into a single entity (as even the U.S. and U.K. do not have a single organization responsible for all intelligence functions); rather, it implies a system that enables a smooth flow of information, as well as accurate and rapid decision-making and execution when conducting ACD.

5. Conclusion

From the perspective of cyber deterrence, the most critical concern that must be pointed out is the issue of interpreting “neutralization,” which forms the backbone of this Japan’s ACD Acts. While this is likely intended to give responsible JMOD and NPA discretion in its application, depending on how it is interpreted, there is a risk that neutralization measures could become meaningless. In the worst-case scenario, it cannot be ruled out

that measures such as requesting internet and/or management service providers to suspend accounts engaged in malicious activities with their consent—which are no different from measures taken prior to Japan's ACD Acts—might be labeled as neutralization. In such a scenario, contrary to the examples in the U.S. and the U.K., not only would there be no effect of deterrence by punishment against attackers who have no reason to fear retaliation, but it would also be impossible to stop attacks utilizing virtually unscathed attack networks, making it highly likely that system enhancements would be undermined from the perspective of deterrence by denial. Within Japan's legal framework, there is always room for debate regarding the establishment of an organizational framework in its final form or its gradual development. It is presumed that the latter approach is taken with regard to ACD as well. However, the imminent threat of cyberattacks necessitates immediate response. While the implementation of additional government reorganization and legal reforms to enable neutralization capabilities that could function as a deterrent is imperative, time-consuming initiatives should be delegated to future legal reforms. Given the impending enactment of the new acts in November 2026, it is essential to formulate pragmatic measures to facilitate effective implementation under the Japan's ACD Acts through amendments to cabinet orders and other pertinent regulations at the time of enactment. It is hoped that this paper will serve as a reference for Japan's efforts.

Footnotes

[1] The Department of Defense, *The Department of Defense Strategy for Operating in Cyberspace, July 2011,* p. 5. https://csrc.nist.gov/CSRC/Projects/ISPAB/documents/DOD-Strategy-for-Operating-in-Cyberspace.pdf (Accessed 27 August 2024)

[2] World Economic Forum, *The Global Risks Report 2017, 12th Edition* (Geneva: World Economic Forum, 2017), p. 11.

[3] JCS, *Joint Planning, Joint Publication 5-0 (JP 5-0)*, 16 June 2017, p. F-4.

[4] J. S. Nye Jr., "Deterrence and Dissuasion in Cyberspace," *International Security,* Vol. 41, No. 3 (Winter 2016/17), 2017, pp. 45-53.

[5] A. Brantly, "The Cyber Deterrence Problem," in *10th International Conference on Cyber Conflict. Proceedings 2018* (Tallinn: NATO CCD COE Publications, 2018), p. 38.

[6] Takahisa Kawaguchi, "Transformation of US Cyber Deterrence Policy: Attribution and Resilience," *Keio SFC Journal*, Vol. 15, No. 2, 2015, p. 84.

[7] E. Iasiello, "Is Deterrence an Illusory Course of Actions?" *Journal of Strategic Security* Vol. 7, No. 1, Spring 2014, pp. 1-2.

[8] W. Lynn, *Remarks on Cyber at the RSA Conference*, San Francisco, California, February 15, 2011. https://www.youtube.com/watch?v=T1PbBy6Gsnc (Accessed 12 June 2024)

[9] K. Giles, "Russia's Public Stance on Cyberspace Issues," in C. Czosseck, R. Ottis, K. Ziolkowski, eds., 2012 4th International Conference on Cyber Conflict (Tallinn, NATO CCD COE Publications, 2012), pp. 64-65.

[10] T. Rid, B. Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies*, Vol. 38, No. 1-2, 2015, p. 30.

[11] John Boyd, "A Discourse on Winning and Losing," https://www.colonelboyd.com/boydswork (Accessed 22 August 2024).

[12] An organisation that handles computer security incidents such as unauthorised access or vulnerabilities on computers.

[13] E. Hutchins, M. Cloppert, R. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," *6th International Conference on Information Warfare and Security*, ICIW 2011, p. 4.
[14] TOR: Abbreviation for The Onion Router, a technology used to anonymise routes on the internet.
[15] Botnets: A network composed of numerous zombie computers hijacked by cybercriminals using Trojan horses or other malicious programs.
[16] For a scale to evaluate possible responses by a state to a cyber incident, see Jason Healey's 2012 framework, which classifies state responsibility for a cyber incident into 10 levels (from ignoring it up to retaliation), depending on the severity. Jason Healey, "Beyond Attribution: Seeking National Responsibility for Cyber Attacks," *Atlantic Council*, January 2012, http://www.acus.org/files/publication_pdfs/403/022212_ACUS_NatlResponsibilityCyber.PDF (Accessed 18 June 2024)
[17] J. E. McGhee, "Liberating Cyber Offense," *Strategic Studies Quarterly*, Vol. 10, No. 4 (Winter 2016), pp. 47-49.
[18] T. McKenzie, "Is Cyber Deterrence Possible?" p. 7. https://media.defense.gov/2017/Nov/20/2001846608/-1/-1/0/CPP_0004_MCKENZIE_CYBER_DETERRENCE.PDF (Accessed 28 August 2024)
[19] Stefan Soesanto, "Cyber Deterrence Revisited," *Air University Press*, October 2021, Vol. 31, No. 2, 2022, pp. 289-324.
[20] F. Egloff, M. Smeets, "Publicly attributing cyber attacks: a framework," *Journal of Strategic Studies*, Vol. 46(3), 2023, pp. 502-533., p. 504.
[21] Nye, op. cit., p. 52.
[22] These refer to the ransomware viruses developed by North Korea and Russia, as well as related incidents. https://www.spf.org/iina/articles/kawaguchi_02.html (Accessed 18 August 2024)
[23] F. Egloff, "Public attribution of cyber intrusions", *Journal of Cybersecurity 2020*, vol. 00, No. 0, p. 9.
[24] JCS, Cyberspace Operations, *Joint Publication 3-12* (JP 3-12), p. xi.
[25] Harold Hongju Koh, "International Law in Cyberspace," *Harvard International Law Journal*, Vol. 54., Dec 2012, pp. 3-4. https://www.academia.edu/73674514/International_Law_in_Cyberspace (Accessed 27 August 2024)
[26] The Department of Defense, "The Department of Defense Strategy for Operating in Cyberspace, July 2011," p. 5. https://csrc.nist.gov/CSRC/media/Projects/ISPAB/documents/DOD-Strategy-for-Operating-in-Cyberspace.pdf (Accessed 24 August 2024)
[27] Takahisa Kawaguchi, "Current Status and Challenges of Cybersecurity in National Security," *in New Issues for the US–Japan Alliance in the Global Commons (Cyber, Space, and the Arctic),* The Japan Institute of International Affairs, 2014, p. 18.
[28] M. Libicki, Cyberdeterrence and Cyberwar (Santa Monica: RAND Corporation, 2009), p. 33.
[29] The Department of Defense, The Department of Defense Cyber Strategy 2018, pp. 1-3. https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBERSTRATEGY SUMMARY FINAL.PDF (Accessed 27 August 2024)
[30] U.S. Cyber Command News, 7 March 2023, https://www.cybercom.mil/Media/News/Article/3320195/2023-posture-statement-of-general-paul-m-nakasone/ (Accessed 27 August 2024)
[31] Cyber National Mission Force Public Affairs, "Before the Invasion: Hunt Forward Operations in Ukraine," 28 November 2022, https://nsarchive.gwu.edu/sites/default/files/documents/rmsj3h-751x3/2022-11-28-CNMF-Before-the-Invasion-Hunt-Forward-Operations-in-Ukraine.pdf (Accessed 18 August 2024)

[32] Cyber National Mission Force Public Affairs, "U.S. conducts first Hunt Forward Operation in Lithuania," 4 May 2022, https://www.cybercom.mil/Media/News/Article/3505610/us-conducts-first-hunt-forward-operation-in-lithuania/ (Accessed 18 August 2024)
[33] Sean Lyngaas, "US military's hacking unit publicly acknowledges taking offensive action to disrupt ransomware operations," https://edition.cnn.com/2021/12/05/politics/us-cyber-command-disrupt-ransomware-operations/index.html (accessed 18 August 2024).
[34] Sean Lyngaas, "US confirms military hackers have conducted cyber operations in support of Ukraine," https://edition.cnn.com/2022/06/02/politics/us-hackers-ukraine-support/index.html (Accessed 18 August 2024)
[35] Harvard University Belfer Center, *National Cyber Power Ranking 2020*.
[36] Richard Norton-Taylor, "GCHQ steps up strategy to combat cyber-attacks, Brown announces," The Guardian, 25 June 2009
[37] Government of the United Kingdom, National Cyber Security Strategy 2016-2021, p. 9. https://assets.publishing.service.gov.uk/media/5a81914de5274a2e8ab54ae9/national_cyber_security_strategy_2016.pdf (Accessed 27 August 2024)
[38] Naoya Yoshino, "'Active Cyber Defense' Secures Public Safety, says NCSC Director," Nikkei-Newspaper, 11 July 2024.
[39] NCSC, https://www.ncsc.gov.uk/information/takedown-service (Accessed 16 August 2024).
[40] Danny Palmer, "GCHQ Chief Reveals Major Cyber Attack Against ISIS—Also Specifically Accuses Russia," CNET Japan, 13 April 2018. https://japan.cnet.com/article/35117687/ (Accessed 13 August 2024)
[41] Masahiro Kurosaki, "The International Legal Framework for Active Cyber Defense—Possibility of Justifying Activities Below Armed Conflict Threshold," *International Affairs*, No. 716, The Japan Institute of International Affairs, 2023, p. 29.
[42] Kazuhiro Nakatani, Keiko Kono, Masahiro Kurosaki, *International Law on Cyber Attacks: Commentary on the Tallinn Manual 2.0*, Shinzansha, 2023, p. 86.
[43] "National Security Strategy," December 16, 2022 (Decision of the National Security Council and Cabinet Decision), p. 30.
[44] The legal titles are the "Act on the Prevention of Damage Caused by Unauthorized Acts Against Important Computers (Act No. 42 of 2025)" and the "Act on the Maintenance of Related Laws Accompanying the Enforcement of the Act on the Prevention of Damage Caused by Unauthorized Acts Against Important Computers (Act No. 43 of 2025)." Their common names are the "Cyber Response Capability Enhancement Act" and the "Related Maintenance Act," respectively. In this paper, these are collectively referred to as 'Japan's ACD Acts'.
[45] As defined in Article 81-3 of the Self-Defense Force Act.
[46] As defined in Article 2, Paragraph 6 of the 'Cyber Response Capability Enhancement Act.'
[47] When implementing neutralization measures against servers or other infrastructure located abroad, even if such actions constitute an violation of sovereignty, they shall be carried out within the limits permitted under international customary law, such as by invoking legal doctrines like the "state of necessity."
[48] As defined in Article 2, Paragraph 3 of the 'Cyber Response Capability Enhancement Act.'
[49] Kurosaki, op. cit., p. 30.
[50] Takashi Seto, "State Cyber Attacks and Public Attribution," NIDS Commentary No. 179, 2021, pp. 4–8.