Expanding the Dyadic Cyber Incident and Campaign Dataset (DCID): Cyber Conflict from 2000 to 2020[1]

Dr. Ryan C. Maness
Dr. Brandon Valeriano
Dr. Kathryn Hedgecock
Jose M. Macias
Dr. Benjamin Jensen

ABSTRACT

This article provides an overview of updates to the Dyadic Cyber Incident and Campaign Dataset (DCID). Whereas most efforts to catalogue cyber incidents focus on curated lists and attack typologies, the DCID uses a standardized set of coding procedures consistent with best practices in social science. As a result, the analysis reveals there is a tendency to exaggerate the use and impact of cyber operations, obscuring their role as an instrument of disruption, espionage, and sabotage, and complements to larger coercive campaigns. The article outlines the construction of version 2.0, which documents rival, state-to-state use of cyber operations as an instrument of power. The expanded dataset introduces additional incidents based on various web-searching methods and human coder cross-validation while also adding new variables for ransomware, supply chain attacks, and connections to ongoing information operations. DCID 2.0 contains 429 incidents representing a critical attempt to scope the domain of conflict among strategic rivals.

INTRODUCTION

In the 21st century, cybersecurity is an increasingly critical issue for competition and conflict among all actors in the international system. The cyber domain,[2] which encompasses digital competition across the physical, logical, and persona layers of cyberspace is not only a site of contestation but a focal point for debates about strategy, defense spending, and alignment of human capital to national security priorities.[3] It is now common to argue that cybersecurity is a top tier security threat that will dominate future battles through the speed of interaction and the fast pace of technological advancement.[4]

Ryan C. Maness (Ph.D., University of Illinois, Chicago, 2013) serves as an assistant professor at the Naval Postgraduate School (NPS), where he oversees the curriculum on information strategy and political warfare. He also directs the DoD Information Strategy Research Center at NPS, which will host its third annual Symposium on Information Strategy and Political Warfare in June 2023. Maness's research includes a state-initiated cyber incidents project, the Dyadic Cyber Incident and Campaign Dataset (DCID) 2.0, which tracks incidents from 2000-2020. He also studies escalation in cyberspace and cyber operations' role in the Russo-Ukraine war. Currently under contract for a book on Modern Hybrid Warfare, Maness has also published works on cyber strategy and Russia's coercive diplomacy. He teaches courses on cyberspace conflict, computer network attack and defense, and conflict in Europe at NPS.

Due to the increasing connectivity of the modern world, competition and conflict in information space are the new norm and international interactions increasingly take place in and through the cyber domain. From digital thefts that sustain authoritarian regimes, to attempts to destabilize the foundations of democracy, digital connectivity provides new options for state and non-state actors to engage in contentious politics. Cyber connectivity is thus a risk as much as it is an opportunity. Since security was not built into these networks, and still is often an afterthought, this pattern of contestation is likely to continue. Therefore, researchers owe the public and policymakers new datasets that identify patterns and trends defining how major state actors use cyber operations against their rivals. This article suggests one set of data collection variables that attempt to address this issue.

Constant threats from adversaries hasten the need for clear, open-source, and timely data on cyber incidents to counter and defend rival cyber operations, which are becoming an ever-growing threat to global stability and connectivity. Cybersecurity is an operational domain of conflict functioning almost wholly without data that might illuminate observers on the scope of the issue. The field has no clear awareness of the baselines for cybersecurity incidents and few methods of collecting information to rectify this problem. This challenge was noted in the policy community, culminating in the Cyberspace Solarium Commission's failed recommendation of a Bureau of Cyber Statistics to collect cybersecurity data.[5]

This article details the expansion of the Dyadic Cyber Incident and Dispute (DCID) Dataset first produced by Valeriano and Maness[6] and expanded into version 1.5 with Maness, Valeriano, and Jensen.[7] Version 2.0 represents a new statement of comprehensive data coding for state-to-state based cyber conflict. With this data expansion, we add additional incidents, expand the timeline, and add in new variables for such factors

Brandon Valeriano (PhD) recently joined Seton Hall University's School of Diplomacy and the Royal Danish Defence College. He also serves as a Distinguished Senior Fellow at the Marine Corps University and Senior Advisor to the Cyberspace Solarium Commission 2.0. He has published six books and dozens of articles on cyber security and international security. Dr. Valeriano has provided testimony on cyber conflict for the United States Senate, the Canadian Senate, and the Parliament of the United Kingdom, also serving as a Senior Advisor to the Cyberspace Solarium Commission. Ongoing research explores conflict escalation, the space-cyber nexus, and data/metric applications for cyber security. Dr. Valeriano is the Area Editor in International Relations for the *Journal of Cybersecurity* and the Series Editor of *Disruptive Technology and International Security* for Oxford University Press.

as ransomware, supply chain attacks, and connections to ongoing information operations. DCID 2.0 has 429 incidents representing a critical attempt to scope the domain of conflict among strategic rivals. An exhaustive search of other data sources has been consulted to identify any missing cases ensuring that this data supersedes all previous efforts to catalog the domain.

In this article, we put cyber data in context, explain our design choices, and outline our early findings. We then review our coding procedures and report reliability statistics. Finally, we report new results on the number and impact of cyber operations over time, demonstrating that cyber operations are on the rise, used differently by each state based on their interests, and that the U.S., its allies, and partners can develop more coherent policies by understanding how and when their adversaries are using the cyber interactions maliciously. Additionally, the best course of action to deter and discourage malicious cyber actions is suggested.

Cybersecurity Data

Current literature and research provide an incomplete picture of the cyber landscape. Organizations that collect cyber event data often encounter extreme roadblocks since cybersecurity operates as a covert or clandestine instrument of power – often undeclared and unattributed.[8] Limited efforts to provide data exist because automated coding of reports of incidents, including using natural language processing methods are complicated by the lack of available non-proprietary data. Transparency of incidents due to monetary or reputation costs also leaves many incidents incomplete or undisclosed in the public record, thereby, complicating data collection.

Four types of data exist in the literature. First, event lists provide information to researchers. Second, annual documents by cybersecurity companies report summaries of information and statistics of known cyber events. Third, while not exclusively data, the wider

Kathryn Hedgecock is an Assistant Professor of International Affairs at the United States Military Academy in West Point, NY. Major Hedgecock's research interests include deterrence and attribution of state-sponsored cyber operations, the impact of emerging technology on the future of warfare, the relationship of public and private entities in cyber defense, and the public response to non-conventional operations short-of-force. She holds a Ph.D. and M.A. from Stanford University in Political Science, a M.Sc. from the University of Oxford in Russian and East European Studies, and a B.S. from the United States Military Academy.

body of literature uses case studies, anecdotes, and analogy to inform strategy. Finally, attack typologies that classify attack types as well as techniques and procedures to support real-time cyber defense operations are included.

For the first category, various organizations have sought to provide data in the form of lists to the community. The Council on Foreign Relations (CFR) Cyber Operations Tracker and the Center for Strategic and International Studies (CSIS) Significant Cyber Incidents provide short summaries of available incidents.[9] The Carnegie Institute also produces a smaller list of attacks involving financial institutions.[10]

These lists lack the ability to provide comparative variables, primarily categorical but also interval, because they do not limit incidents relative to analysis.[11] Instead, these lists serve as a qualitative summary of incidents according to their respective criteria but not formatted in a manner consistent with the scientific study of conflicts.[12] Furthermore, no summary statistics or quantitative products are provided alongside, limiting their true potential to social science. Without awareness of the norms and process of dataset collection, these sources fail to build what is typically thought of as a dataset, a well-planned effort to catalog existing efforts in a community including associated variables and information to support replication.[13] What is included, and excluded, plus the reliability checks of the data are missing, leaving the community largely in the dark without valid sources of data.

A graphical representation (Figure 1) outlines the types of data available to researchers. CSIS and CFR code known incidents regardless of target and attacker. Comparing all known incidents from one time to another provides little insight for comparative analysis due to unclear selection methods. For the specific purpose of measuring the political impact of attributed incidents and testing of theories on escalation, deterrence, and

Jose M. Macias is a Master of Public Policy (M.P.P.) candidate at The University of Chicago's Harris School of Public Policy and a Fellow with The Pearson Institute for the Study and Resolution of Global Conflict. Born and raised in Calexico, CA, he's a first-generation Chicano graduate with dual B.A.s in International Relations and Political Science from UC Davis. Jose has an extensive background in national security studies and cyber conflict research, with professional experiences with the U.S. Department of Defense, Army Cyber Command, U.S. Congress and the Center for Strategic and International Studies (CSIS). Now at Harris, he serves as a Teaching Assistant for R programming and co-leads a cyber incident coding project at the Naval Postgraduate School with Dr. Ryan C. Maness and Dr. Brandon Valeriano.

persistent engagement, DCID collects incidents with reasonable attribution of responsibility and evidence in the public record. Additionally, since the purpose is to capture activity of contentious states, DCID collects only incidents involving nation-state historic rivalries.[14]

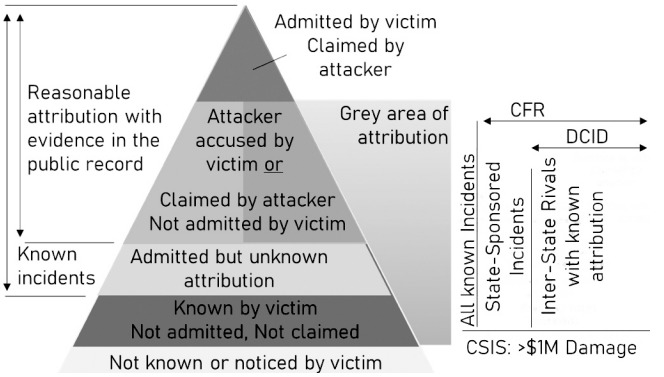

Figure 1: Types of Cyber Incidents

Second, cybersecurity companies produce several annual reports with claims of increasing quantity and severity of attacks, all released for a significant fee. Yet, these statistics limit the scope to the clients of their software or services. Little comprehensive data without significant bias exists to provide statistics that can withstand academic scrutiny, with compilations of secondary and tertiary sources crowding out primary sources and surveys in these reports.[15] Additionally, cybersecurity companies do not reveal or share proprietary data, making compilation, fact checking, or comparison of statistics difficult to impossible.

Third, the quality of research methods across cyber strategy literature varies significantly, with several well-researched case studies. However, much of the literature relies on logic and analogy, rather than facts. The existing academic debates on cyber strategy continue to diverge in their conclusions rather than converge, with deterrence advocates and persistent engagement advocates becoming more entrenched in their convictions.[16] This divergence occurs because little to no convincing data exists to persuade either side of what the

Dr. Benjamin Jensen is a Senior Fellow at CSIS for future war, gaming, and strategy in the International Security Program and is also a Professor at Marine Corps University School of Advanced Warfighting. Over the past decade, Dr. Jensen has shaped defense analysis, crisis response, and military planning through his partnerships with numerous military, government, and international bodies, including DARPA, NATO, and the U.S. Army. Author of four influential books, his most recent work centers on military innovation and AI. He served as the Senior Research Director and lead author for the U.S. Cyberspace Solarium Commission. Dr. Jensen holds a PhD from the American University School of International Service and continues to serve as a reserve officer in the U.S. Army.

facts are to uphold any claims of one side or criticism of the other.

Attack typologies can be useful for defense but lack systematic comparisons, and do not help establish patterns and trends outside of attack types. It is therefore a different level of analysis, where it is not about statecraft and more about technical dimensions of best practices in terms of technical defense. One attack framework created by MITRE Corporation named ATT&CK, utilizes a mixed method of technical and behavioral collection methods, and could be incorporated into our collection processes in subsequent versions of DCID.[17]

Overall, there are multiple types of social science analyses available to researchers, but the genre of quantitative social science, is largely missing from current analysis, thereby revealing the precarious foundations of the field of cybersecurity.[18] As outlined in previous versions of DCID, a codebook exists to ensure researchers understand how the data was coded.[19] Previous publications outline the data collection fields that include the countries involved, the date, website sources, and ten additional categorical classifications. New to this dataset are binary indictors for supply chain, ransomware, and information operations, as well as a categorical variable on the infrastructure sector, as outlined by the NIST cybersecurity framework.[20]

The peace science data collection revolution brings us to the methods utilized for the collection of the data contained in DCID. Two datasets, the Militarized Interstate Dispute Data (MID) and International Crisis Behavior Data (ICB), are efforts that contributed to this revolution.[21] Over the decades, researchers in the international relations field moved beyond narrative descriptions absent structured, coding methodologies that lent themselves to replication studies and searching for aggregate patterns. In the case of Cold War era pioneering data efforts, the use of early punch card computing and coding schemes was designed to under-

stand the causes of war and which crises were more likely to escalate. It was academics trying to serve the broader policy community. Here, we make a similar intervention, trying to take an inductive approach to knowledge construction that facilitates bridging the gap and offering empirical data other researchers can use and policymakers can review to develop cyber policies and strategies.[22]

The state of data collection leaves cybersecurity, national security, and social science research in a vacuum. Those researching trends in cyber incidents are left without clear sources of unbiased, peer-reviewed information. These pathologies provide the motivation for this project as we bring the DCID data into the new decade.

Coding Cyber Incidents

To code cyber incidents means isolating in time and space given available information regarding attempts by rival states to launch a major cyber operation in pursuit of foreign policy objectives. This definition sets the term "incident" apart from everyday cybercrime and constitutes a range of activities from leveraging patriotic hackers to high-end cyber tools exploiting configuration vulnerabilities in national networks. Each incident might include hundreds if not thousands of individual attempts to breach a network and deposit malware – just as a major military operation encompasses countless tactical engagements and battles. In the same way that earlier work on militarized disputes and crises isolated particular acute state interactions, the DCID attempts to isolate how states compete in and through cyberspace.

As the United States (U.S.) struggles to implement breach data notification reporting laws that would provide data for awareness of cyber incidents, a series of consequential cyber operations occurred from 2020 to 2021.[23] From SolarWinds to the Microsoft Exchange vulnerability to the Colonial Pipeline ransomware attack, the U.S. is reacting to incidents with little comprehensive awareness of trends and patterns. Data are necessary to anticipate future challenges and to measure the effectiveness of our current strategy.

There are three reasons the national cyber strategists require a publicly available database of cyber operations. First, there is the issue of strategic planning and budget justification. In the U.S., DoD must demonstrate that it is meeting the requirements for cyber forces at the strategic level. Publicly available data can provide an open-source method for general analysis that demonstrates the requirements for offensive and defensive forces. The rationale for funding and resources does not withstand public scrutiny without a medium to communicate trends in cyber operations.

Second, publicly available data provide an indication of what is known for larger audiences, and not just military practitioners. Data that engage public sources can be leveraged to understand how cyber operations become known to the public, from what sources, and for what purpose. We also have no reason to assume that what is revealed publicly is not a representative sample of the presumably larger, covert operations.

Finally, and perhaps most importantly, the international community is interested in the sponsorship of public goods that allow the academic community to participate and contribute to scholarship that advances the theoretical aspects of cyber conflict theory. As such, our dataset is a public good that is free to access and use for the academic community. Academics debate persistent engagement, deterrence, coercion, and escalation, but do so with little hard data to demonstrate the efficacy of each of their theories.[24] Without data, evidence on the competition of theories is absent, and new theories will not mature, which leads to suboptimal policy.

We manage the collection of the data in a novel manner, using a combination of academics, U.S. military staff officers, and student interns. For this research, a cadre of researchers provided guidance to twelve student interns in the Virtual Student Federal Service (VSFS).[25] Using an interactive process, the team developed a method of collecting, reviewing, and filling out the data needed to create a full and original dataset that included a plethora of independent variables and sourcing information.

After an introduction to the existing academic literature, including journals, books, blogs, websites, and prior data collection efforts, the students used publicly available information (PAI) to expand on the DCID 1.5 dataset, which gathered data through 2016.[26] For inclusion of cyber incidents in this dataset, there must be open-source evidence that a nation-state is culpable for the cyber operation in question.[27] After an initial round of data gathering, the students cross-reviewed the work of other students. To ensure data quality, a team of experienced researchers reviewed the students' work. At the end of their efforts, reliability statistics were collected after a review of the final coding outcomes.

As outlined in previous versions of DCID, the codebook exists to ensure researchers collect data accurately.[28] Previous publications outline the data collection fields, which include the countries involved, the date, website sources, and ten additional categorical classifications. New to this dataset are binary indicators for supply chain, ransomware, and information operations, as well as a categorical variable on the infrastructure sector, as outlined by the NIST cybersecurity framework.[29] Since the researchers are primarily English speakers, we accept the Anglophone bias inherent in our research methods and acknowledge this caveat.

Due to the covert or clandestine nature of cyber operations, as well as the limited public reporting on state-initiated action in the domain, the coding processes for this project are human-driven. Although machine learning web-scraping processes have been proposed,[30] we concluded that the nuance and meticulous procedures needed for coding cyber operations require a human touch. Training processes include the instruction of uniform search terms and procedures by project leaders so that accuracy and replicability are assured. Project leaders then verify all incidents coded for final inclusion or exclusion into the dataset.

Researchers should be aware of biases and caveat results when conducting analysis using the DCID. This issue is common across all datasets; here we provide an open-source dataset that others can use and modify to suit their needs. We also review all other publicly known incidents

coded in other datasets to ensure complete coverage. Many incidents were discarded due to incomplete sourcing, the likelihood that are criminal operations and not state directed, and the bias to report incidents that attack financial systems. For example, some datasets include an overabundance of cryptocurrency attacks that we could not attribute to state action, or even external attacks, since many operations are insider attacks.

Dataset Details

The core of any dataset is the explicit original variables identified and documented. In DCID 2.0, there are several variables coded for each cyber incident. The first are the start dates and end dates for each incident, reported as accurately as possible. For the start date, most reports analyzed gave approximate month and specific year. The end date given for each operation is the approximate earliest public reporting of the termination of the incident.

While multiple terms for actions in cyberspace exist across the academic databases, the term "incident" has a very specific meaning in U.S. code,[31] which we build on here.[32] "The term "incident" means an occurrence that (a) actually or imminently jeopardizes, without lawful authority, the integrity, confidentiality, or availability of information or an information system, or (b) constitutes a violation or imminent threat of violation of law, security, policies, security procedures, or acceptable use policies."[33] We generally avoid the term cyber-attack since it is contested and devoid of meaning in popular usage.[34]

Next is the cyber method variable, which consists of four overall categories. The first is Vandalism (Website defacements, propaganda) where hackers use SQL injection or cross-site scripting (forms of command code) to deface or destroy victims' web pages. Although rather benign, these attacks may have important psychological effects on their intended audience. The second category is denial of service (DDoS, Botnets, data blocking): DDoS attacks flood particular Internet sites, servers, or routers with more requests for data than the site can respond to or process. Such attacks are coordinated through "botnets," or a network of computers that have been forced to operate on the commands of an unauthorized remote user. The primary impact of DDoS attacks via botnets is the temporary disruption of service. The third category is network intrusions ("trapdoors" or "trojans" and backdoors).[35] Spear phishing is utilized to inject these cyber methods into networks. Here the initiator sends emails to employees or contractors of the targeted network and, if the email is opened, the intrusion is introduced to the system. Finally, the fourth category is network infiltration with examples of attacks including logic bombs (wiper malware), viruses, worms, and keystroke logging.[36] These methods force computers or networks to undertake tasks they would normally not undertake.

Target type is the next variable coded and comes in three forms. The first is private/non-state includes targets such as the financial sector, power grid, or multinational corporation (MNC). The second is government non-military variable, with the U.S. State Department and Department of Homeland Security (DHS) government websites, or the Office of Personnel Management

(OPM) as primary examples. Last is the government military variable, where entities such as DoD, U.S. Cyber Command, or U.S. Strategic Command are primary examples here.

The next category for the cyber variables coded in DCID is titled "strategic objective for initiators" and comes in four forms: 1) Disruption: Examples include taking down websites or disrupting online activities. These are usually low-cost, low-pain incidents such as vandalism or DDoS techniques. 2) Short-term espionage: These are acts that gain access which enables a state to leverage critical information for an immediate advantage. 3) Long-term espionage is an event that seeks to manipulate the decision-calculus of the opposition far into the future through leveraging information gathered during cyber operations to enhance credibility and capability, an example being China's theft of Lockheed Martin's F-35 plans. Finally, 4) Degrade is the category that includes attempts of physical degradation of a target's capabilities.

Our next variable is binary and labeled "cyber-enabled information operation" (CIO). For our purposes, we want to distinguish information warfare operations from the more universal awareness of attacks in and through cyberspace as conceived by the public. We also want to avoid the broad definition of information operations as defined by the U.S. military, which includes electronic warfare, psychological operations, and social network exploitation, among others. To avoid muddying the waters more, we instead focus on information operations in the midst of cyber incidents, what we call cyber-enabled information operations (CIO).[37]

The restriction to operations launched during ongoing cyber operations allows us to focus on how information and data are weaponized adjunct to cyber incidents and campaigns. Rather than documenting more broad attacks on truth and meaning, we seek to understand how information can be manipulated and utilized to message in a psychological operation against a target. Its plausible that CIO's are improved using demographic or personal data stolen from initial cyber incidents, to personalize or tailor a specific message to manipulate a specific audience.

The next variable is titled "objective achievement" and is binary: Did the cyber incident achieve its intended purpose is the main question asked of coders. For example, did the disruptive attack successfully shut down a website via denial of service? Did an espionage technique breach the intended network and steal the information it sought to acquire? Did the degradation achieve damage to its intended target?

The "Concession" variable is where a binary score of the presence or absence of a concessionary behavioral change. Coding behavior change can be considered quite subjective, and therefore collaboration with multiple peers and researchers has been utilized in the coding of this variable. If there is an observable change in foreign policy by the target state as a result of the initiator's cyber operation, we code the presence of a concession. These events are quite rare, as it has been found that there are limitations on the coercive power of cyber operations.[38]

In terms of the severity of the impact of each cyber incident, a revised severity scale is included in this version of DCID. The scale is interval and is numbered 1 through 10, with "1" indicating

the least severe cyber operation and "10" denoting the most severe possible incident. It must be noted that in the DCID version 2.0, no state-initiated cyber operation has evoked a severity score above "6."[39] For a cyber incident to be coded with a severity score, there must be a successful breach, where widespread destruction has been found on either a single or multiple networks. An example of an incident with this score is the Iranian degradation attack on Saudi Arabia's oil giant Aramco, in which wiper malware erased the hard drives on nearly 30,000 workstations in 2012.

Cyber incidents coded with severity scores 1 through 5 have traits as follows: an incident coded as "1" are the more passive cyber operations, where packet sniffing and probing, as well as spyware that has not been activated and is found before it has been, are examples of these low-level operations. Incident levels labeled as "2" are forms of cyber harassment, propaganda, and events that deny access to information or disrupt daily activity on a network for a short duration. Incidents that are coded with a severity score of "3" are most espionage campaigns, in which we see the theft of valuable or classified information from a single network. Cyber incidents that are given a "4" severity score are coded as such when there are multiple network and widespread data breaches and espionage activity from the initiating state. Finally, incidents receiving a "5" severity score are those that infiltrate either a single or multiple critical networks and attempt physical destruction.

Next is "Damage Type" conceptualized from scholars and is categorical ranging from 1 to 4.[40] Where (1) refers to "Direct and immediate" damage and costs that are felt immediately. Then (2) "Direct and delayed" damage by an incident that takes months if not years to disrupt or damage. Next is (3) "Indirect and immediate" damage which was not the original intent of the imitator. Examples include reputational damage or loss of confidence in the target by an audience. Lastly, (4) "Indirect and delayed" damage and costs for incidents that would be felt at a future point in time.

Finally, the last variable carried from DCID 1.5 is "Specific political objective." The aim is deciphering why the cyber incident was launched in the first place. For example, for the Sony Hack the objective was to stop the release of the movie *The Interview*. A maximum of two political objectives is allowed.

Considering the changing trend of cyber incidents and to improve analysis, the authors updated DCID with three new variables. The first variable is categorical and labeled "Critical Infrastructure" (CI). This additional variable differentiates between targets vital to national and economic security of the nation as designated by DHS and the Cybersecurity and Infrastructure Security Agency (CISA).[41] Further, CI is coded 1-16 to represent the 16 sectors. In addition, noticing a blind spot, the authors added "17" to represent academic institutions that produce critical and innovative research for the nation.[42] The categorization of entities will help policymakers and industry alike understand the evolution of targets as rivals shift interest and seek influence across sectors.

Our second addition is a binary variable labeled "Supply Chain." We consider targeted companies that serve or provide a service to any of the critical infrastructures listed (1-17). Further, the service or product to the critical infrastructure may be in different forms such as software for network, technology, or IT management. For example, it may include billing systems, cloud services, or remote access systems like a virtual private network (VPN). SolarWinds serves as a reminder that cyber actors target supply chain firms and attempt to manipulate products for backdoors, steal passwords or relevant data, and disrupt or engage in espionage.

The final addition is a binary variable labeled "Ransomware." An incident is coded as ransomware if a target suffers from a loss of data, control, unauthorized encryption, or otherwise locked out of their systems until a "Ransom" is paid in currency, either hard or digital. Previously recorded incidents were retroactively updated with the new variables for a complete set.

In practicality, any set of incidents gathered in a dataset includes only a biased subset of existing cyber actions. The bias is important for international security analysis, as it represents those attacks publicly known between two actors.[43] This provides a semi-complete picture of the landscape for deterrence, persistent engagement, and escalation studies, as well as comparative statistics between time periods. *Despite the limitations of data collection efforts, comprehensive data of what is known is more valuable than speculation on what is unknown.*

The Devil is in the Details – Cyber Relations over Time

Figure 2 shows a graphical representation of cyber incidents categorized by strategic objective for initiators for the years 2000-2020. Espionage campaigns make up over 61 percent of the data, indicating that the battle for information and intelligence is where states are finding utility with cyber operations. Whether it be short-term espionage incidents that aid in operations in the information environment, or long-term ones that steal intellectual property to fill a technological gap in one's military organization, it is information asymmetries and intelligence gathering that provide tangible value to initiating states' national security objectives.

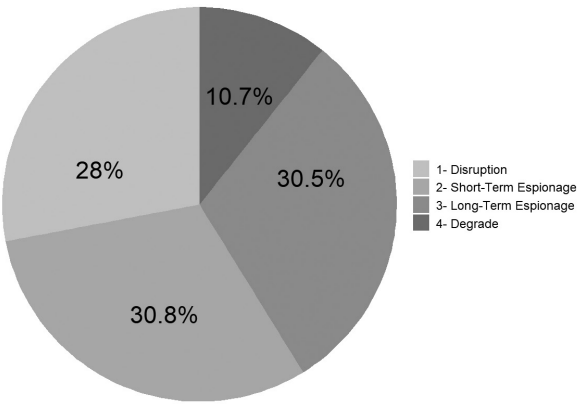

Figure 2: Incidents by Strategic Objective: 2000-2020

Disruptions are utilized roughly 28 percent of the time, are low-risk and inflict pain, but can be interpreted as a type of ambiguous signal intended to demonstrate displeasure with a previous action of the target state.[44] Degradations, utilized just over 10 percent of the time, are riskier and escalation-prone as their intent is to permanently deny the target a capability it once previously

had. These types of cyber incidents are also the most likely to produce concessions from the target state.[45] However, they are expensive as well as risky, and are usually only developed and deployed by the powerful states, making them rarer.

Table 1 shows the strategic objective categories parsed by the initiating country for the previous iteration of DCID that covers 2000-2016. China, Russia, Iran, and North Korea, the four major cyber adversaries of the U.S., lead the world in terms of initiating countries, and target the world's only superpower often. This will be used in Table 2, which provides the updated list of strategic objective initiators.

Table 1. DCID 1.5: 2000-2016

Initiating Country	Disruption	Short-term espionage	Long-term espionage	Degrade	Total
China	17	37	18	2	**74**
Russia	19	23	12	11	**65**
Iran	8	9	12	4	**33**
North Korea	13	3	6	4	**26**
United States	0	1	9	11	**21**
Pakistan	11	1	1	0	**13**
Israel	1	0	7	1	**9**
India	6	1	0	0	**7**
South Korea	4	3	0	0	**7**
Japan	3	0	0	0	**3**
Ukraine	1	0	0	1	**2**
Turkey	1	1	0	0	**2**
Georgia	1	0	0	0	**1**
Syria	1	0	0	0	**1**
Taiwan	0	1	0	0	**1**
Vietnam	0	0	0	1	**1**
Total	86 (32%)	80 (30%)	65 (24%)	35 (13%)	266

Source: Data from Maness, Valeriano, and Jensen, 2019.

Table 2 shows the updated descriptive statistics by strategic objective, with the four authoritarian countries leading the pack once again. Given the fact that most democratic countries around the world are restrained from launching frequent cyber operations, this is not surprising. In the U.S., for example, for a long time the authority to launch an offensive cyber operation (OCO) lay solely with the President. After the revision of the DoD Cyber Strategy in 2018, the commander of USCYBERCOM now has Title 10 authority to launch cyber operations that are considered below the threshold of armed conflict, with the President retaining the sole authority for those considered to be above that threshold.[46] Furthermore, the U.S. does keep most of its cyber methods and actions classified, making this open-source data more than likely undercounting U.S. cyber operations as a result.

Table 2. DCID 2.0 Update: 2000-2020

Initiating Country	Disruption	Short-term espionage	Long-term espionage	Degrade	Total
China	22	49	42	2	**115**
Russia	28	39	30	16	**113**
Iran	13	21	21	10	**65**
North Korea	24	10	19	6	**59**
United States	1	1	9	9	**20**
Pakistan	13	4	3	0	**20**
Israel	2	0	7	1	**10**
India	6	2	0	0	**8**
South Korea	4	3	0	0	**7**
Japan	3	0	0	0	**3**
Ukraine	1	0	0	1	**2**
Turkey	1	1	0	0	**2**
Georgia	0	1	0	1	**2**
Syria	1	0	0	0	**1**
Taiwan	1	0	0	0	**1**
Vietnam	0	1	0	0	**1**
Total	120 (28%)	132 (31%)	131 (31%)	46 (11%)	429

China and Russia still top the list as the most frequent initiators of cyber operations, with their great power adversary, the U.S., being the primary target. China has continued its espionage campaigns, and these have largely increased during the years 2016-2020. After a slight détente with the U.S. in 2015 following the Office of Personnel Management (OPM) hack, it seems that China is back to its old ways. Russia's hacking has also spiked considerably, with the U.S. and Ukraine targeted the most by these recent cyber operations. Iran has increased its cyber presence in the Middle East, and North Korea has turned to cybercrime, including cryptocurrency theft, to bulk up its numbers during these four years.

Table 3 shows the countries that have been victims of cyber operations in descending order of frequency. The U.S., South Korea, Ukraine, India, and Israel have been subject to cyber operations with the greatest frequency. Increased intensity between rivals is the likely culprit behind this increase. The U.S. continues to be barraged by Russia and China; South Korea and North Korea continue their intense rivalry in cyberspace; Ukraine is on the brink of a Russian invasion at the time of this writing; and the India-Pakistan and Israel-Iran rivalries continue to endure. Furthermore, Table 3 indicates the percent increase in cyber incidents each state has experienced between DCID 1.5 and DCID 2.0. While most nations have continued to see increased frequency of operations, the United Kingdom and Turkey stand out as the two countries experiencing the greatest increase in operations, mainly due to low-incident frequency in a database that includes exclusively rival dyads.

Table 3. Countries by Frequency of Target (Comparison of DCID 1.5 & DCID 2.0)

Targeted Country	DCID 1.5	DCID 2.0	Percent Increase
United States	82	138	68
South Korea	26	46	59
Ukraine	15	28	87
India	20	28	40
Israel	11	20	82
Japan	13	18	38
Iran	14	17	21
United Kingdom	3	13	333
Russia	11	12	9
Saudi Arabia	7	12	71
Taiwan	7	12	71
Turkey	4	11	175
Georgia	6	8	33
Pakistan	7	8	14
Vietnam	4	8	100
Philippines	5	8	60
China	7	7	0
Lithuania	4	6	50
Germany	3	5	67
France	3	4	33
Estonia	4	4	0
Canada	2	3	50
Poland	3	3	0
Syria	1	1	0
Lebanon	1	1	0

Figure 3 displays the frequency of cyber incidents in the data over time. The year represents the end date of the operation. DCID 2.0 includes operations as recent as 2021 and subsequently captures additional operations that were not yet public upon release of version 1.5. Since the significant jump in 2014, the prevalence of cyber operations has remained high. However, the increase in initiated cyber operations remains concentrated among authoritarian countries, while democracies continue to be targeted at increased rates, demonstrating the imbalance between these two types of systems.

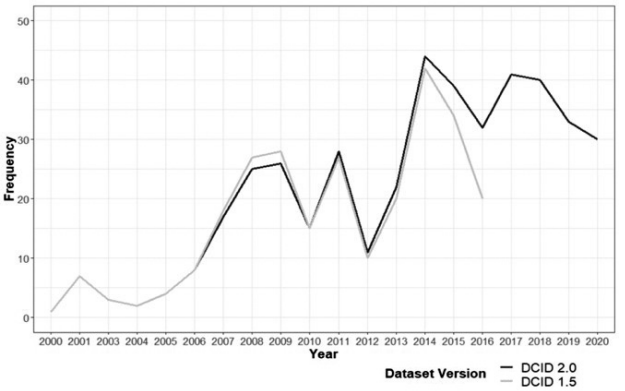

Figure 3: Frequency of Cyber Incidents Over Time (DCID 1.5 vs. DCID 2.0)

A critical question often ignored in cybersecurity is the efficacy of operations. Do cyber operations achieve their objectives when hacking their targets? Other key questions include the participation of third parties or the connection to information operations. Table 4 represents the percentage of cyber operations that have generated concessions, have had multiple initiators, have a cyber-enabled information operation result in the initiation, and have achieved their intended objective in terms of breaching the targeted network. With the addition of new incidents into DCID 2.0, the efficacy of the operations has dropped across all dimensions. There have been no new observable behavioral changes as a direct result of a cyber operation for the years 2016-2020. China has focused on disinformation campaigns along with its more tailored espionage operations, which are arguably less coercive.

In this dyadic dataset, we also see very few joint state initiatives, as countries seemed to keep to themselves when launching their operations. In this version of the data, cyber-enabled information operations cover about a fifth of the dataset. As influence operations and disinformation

continue to be weaponized, we expect this trend to grow as future updates of the data are produced. There is a clear connection between cyber operations and digital influence operations, but the basis for this is unclear.

Table 4. Incident Categorization, Count, and Percent of Incidents

	Total Incidents	Concession	Third Party Initiator	Information Operation	Objective Achievement
DCID 1.5	266	12 (4.5%)	25 (9.4%)	73 (27.4%)	242 (91%)
DCID 2.0	429	12 (2.8%)	32 (7.5%)	96 (22.4%)	368 (85.8%)

Severity and Impact

Figure 4 shows cyber incident severity over time. Although the number of cyber operations continues to grow as states integrate them into their foreign policy and military strategies, the relative severity over time has remained quite constant. Most cyber operations have severity scores between "2" and "4," which ranges from disruptions and propaganda to network espionage in these types of operations we see states find the utility of cyber incidents as force multipliers in hybrid operations. Russia utilizes cyber operations to sow chaos and confusion by spreading propaganda in target countries. China's espionage campaigns, in its view, are contributing to its rise as an economic and military superpower. It is these types of operations we expect to see more in the future as governments learn the primary utility of cyber operations.

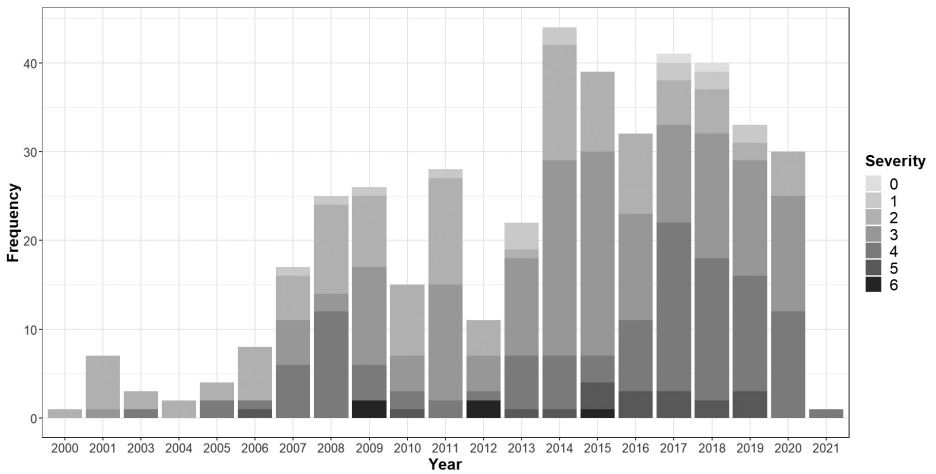

Figure 4. Severity of Incidents Over Time: DCID 2.0

Table 5 shows which severity levels have increased the most between DCID version 1.5 and 2.0. As expected, we see severity levels "3" and "4" rise the most, as these categories encompass primarily espionage campaigns. Attempted physical destruction, which are operations under severity level "5" have also risen sharply. Strategic sabotage is being attempted more often, but with limited success.[47]

In terms of the severity levels between "7" and "10," we have yet to observe these more

dangerous operations. States have yet to leverage cyberspace to take down financial markets, turn out the lights for a long period of time, or intentionally contribute to the loss of life. There are deterrence and restraint mechanisms at play here, as no nation wants to be the first to take these more severe actions, which would likely lead to a more dangerous and unstable cyberspace.[48]

Table 5. Cyber Severity Changes between DCID 1.5 and 2.0

	0	1	2	3	4	5	6
DCID 1.5	0	9	92	97	53	12	3
DCID 2.0	2	15	115	159	115	18	5
Percent Increase	200	67	25	64	117	50	67

Ransomware, Supply Chain, and Critical Infrastructure

Ransomware attacks are attributed to only four states during the years 2000-2020. Looking at Table 6, acts of ransomware remain a motivation of non-state actors, primarily cyber-criminal groups. The Russia-based, non-state group REvil became internationally known due to its role in the Colonial Pipeline incident, which shocked oil supply as well as prices in the Eastern U.S. in 2021.[49] The group has since been disbanded and its members arrested by the Russian government. Russia has launched false ransomware incidents, such as NotPetya, to act as smokescreens so that it can more freely maneuver within Ukrainian government networks, but this method is not common.[50]

Table 6. Frequency of Ransomware Incidents between DCID 1.5 and 2.0

Initiating Country	DCID 1.5	DCID 2.0	Percent Increase
Russia	3	4	33
China	1	3	200
North Ko-rea	0	3	300
Pakistan	1	1	0
Total	5	11	120

Given the potential lucrative earnings that successful ransomware attacks produce, these operations benefit states with limited gross domestic product output and trade. As such, we can expect isolated and desperate actors like North Korea to continue to use ransomware as an alternative, low-cost method to fund its regime. North Korea has grown quite adept in cybercrime in recent years, with the theft of cryptocurrency as well as the famous WannaCry ransomware operation being prime examples. In 2021, the North Korean regime stole nearly $400 million in cryptocurrency.[51] WannaCry, launched in 2017, led to nearly $4 billion in losses globally.[52] Whether to earn currency for the regime or cause global disruptions, North Korea is punching above its weight regarding damaging criminal operations.

Table 7 highlights how cyber incidents targeting supply chains are increasing in recent years. China, Russia, Iran, the U.S., and North Korea are the most active states, with 88 percent of all incidents. Exploiting global supply chains and supply-side trust entities with malware is now a common tool in the cyber domain. The reason is simple, affiliates and trusted entities generally spend less on cyber defenses, focus less on updating their equipment, and are ripe targets for phishing attacks. Attacking the weaker third-party actor is a viable means of penetrating the hardened bunker that is usually state-based cyber entities.

Analyzing supply chain incidents further, Figure 5 visualizes the share of supply chain incidents by Initiator from 2000 to 2020. Through the years, there has been an average of

8 supply chain incidents per year among all initiators. In the first decade (2000-2010), China targeted the highest quantity of supply chain entities with 16 incidents out of 37 for the entire decade, with a sharp drop in the final year (2010). The U.S. follows China with 9 incidents, followed by Russia with 5. However, in the next decade (2011-2020), there is a significant growth in targeting supply chains by Russia with a total of 33 incidents, overtaking China (31), Iran (19), North Korea (11), and the U.S. with only 4 incidents. In all, the first decade of the twenty-first century, supply chains experienced 37 incidents compared to 110 incidents for the second decade. This is a 197% increase and highlights the strategic focus of China, Russia, Iran, U.S., and North Korea.

Table 7. Frequency of Supply Chain Incidents between DCID 1.5 and 2.0

Targeted Country	DCID 1.5	DCID 2.0	Percent Increase
China	32	48	50
Russia	23	38	65
Iran	10	19	90
U.S.	10	13	30
North Korea	5	12	140
Israel	5	6	20
Pakistan	1	4	300
South Korea	2	2	0
Vietnam	1	1	0
India	1	1	0
Taiwan	1	1	0
Turkey	1	2	100
Ukraine	1	1	0
Total	93	148	59

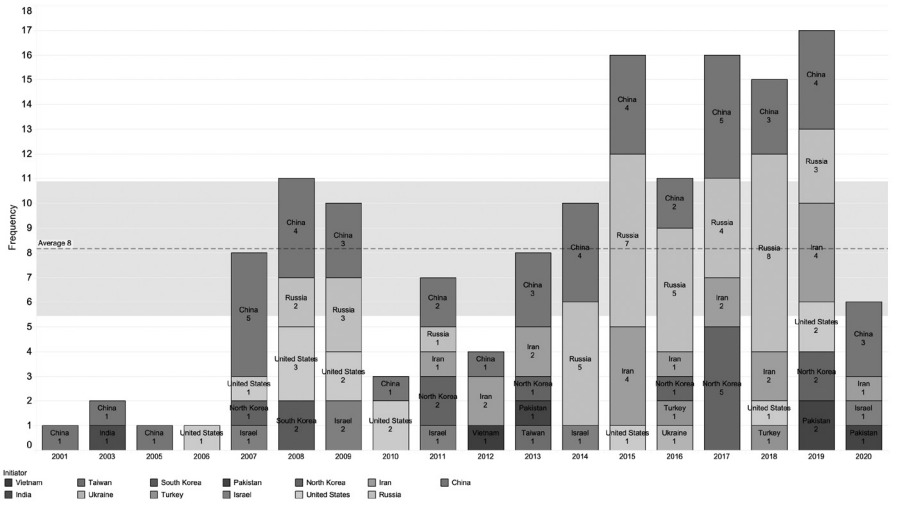

Figure 5. Supply Chain Incidents Over Time DCID 2.0

** Note: The researchers found no supply chain incidents in the years: 2000, 2002, 2004 and are therefore absent from this time-series graph.*

Table 8 breaks down the frequency of attacks on Critical Infrastructure (CI) between DCID 1.5 and 2.0. The most targeted CI among both versions is government facilities, with about 42% of the total share. The next most frequent CI sectors targeted are the academia and information technology sectors, which make up around 10 percent each. With strategic competition among states on the rise, we expect these sectors to continue to be targeted.

Public health facility targeting is on the rise, which is concurrent with the rise in ransomware

Table 8. Frequency of Critical Infrastructure Incidents between DCID 1.5 and 2.0

CI Sector	DCID 1.5	DCID 2.0	Percent Increase
Government Facilities	128	179	40
Academia	18	41	128
Information Technology	29	44	52
Financial Services	18	31	72
Energy	18	29	61
Defense Industrial Base	12	19	58
Public Health	4	17	325
Communications	7	16	129
Transportation	7	14	100
Commercial Facilities	12	12	0
Nuclear Reactors, Materials, Waste	8	13	63
Chemical	2	5	150
Critical Manufacturing	3	4	33
Water and Wastewater Systems	0	3	300
Dams	1	1	0
Emergency Services	0	0	0
Non-CI	1	1	0
TotalS	268	429	60

attacks during the COVID-19 crises during 2020 and 2021. Despite conventional wisdom, there is no significant evidence of increased attacks on nuclear facilities, financial services (economic cyber warfare), or water treatment facilities. This finding does not invalidate the need to protect these sectors, but the idea of a dramatic increase in these sorts of attacks is not evidenced by this version of the data.

To analyze CI further, Figure 6 visualizes CI according to initiating state. Once again, China, Russia, Iran, and North Korea take the lead in targeting CI over the U.S.. Where China has conducted the most operations against CI, it focused primarily on Government Facilities (54), followed by Academia (13) and Information Technology (12).

Russia follows China closely in focusing more on Government Facilities (56), then Academia (10) and Energy (9). Further, this analysis highlights that, while both advanced in cyber capabilities, each has separate aims in its operations. China is focused on cyber operations that steal information from the private and public sectors for diplomatic aims or improving its markets through intellectual property theft. On the other hand, Russia is less interested in improving its domestic market through intellectual property theft and more interested in cyber as a tool of disruption.

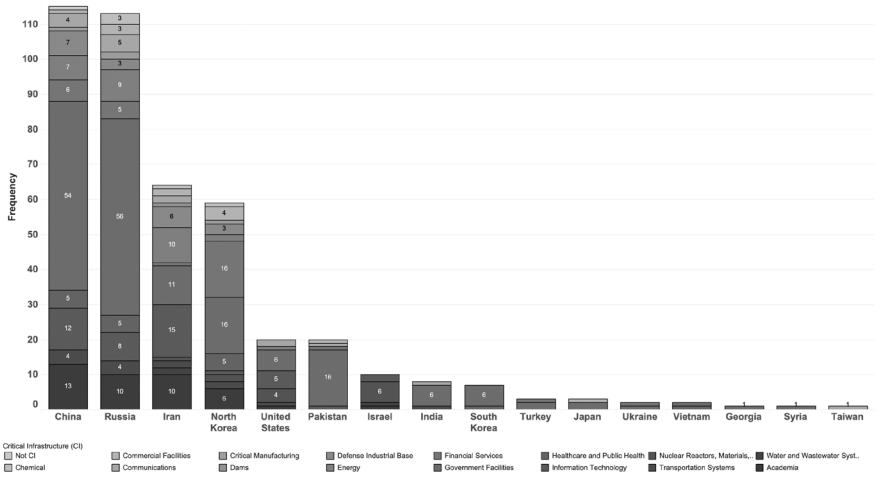

Figure 6. Critical Infrastructure Incidents by Initiator

Next are the nascent but rapidly improving cyber states. Iran focused on Information Technology (15), then Government Facilities (11), and tying for third is Academia and Energy for 10 incidents each. Instead, Iran's cyber operations are tailored less on circumventing international sanctions and toward a diplomatic end. Iran seeks a regional power rebalance by targeting its top rivals, Israel and Saudi Arabia, signaling its resolve.

North Korea pursues a different approach as the other nascent power. This is because its top two targets are Financial Services and Government Facilities, with 16 incidents each, and its third target is Academia (6). North Korea uses cyber operations to circumvent international sanctions to fund their regime, spy on South Korea and academia that influences policymakers. With unsuccessful diplomatic efforts under the Trump administration, North Korea is likely to explore other ways to leverage its cyber operations in seeking sanction relief.

By contrast, as an advanced cyber actor, the U.S. is showing restraint, owning only 5 percent of all CI incidents across 20 years. This might be in part due to the covert nature of U.S. cyber operations, the desire to maintain asymmetrical capabilities a secret, or leadership refraining from declaring operations against rivals. This could also be explained by the lack of technical analysis available by cybersecurity firms when analyzing suspected U.S.-initiated incidents.

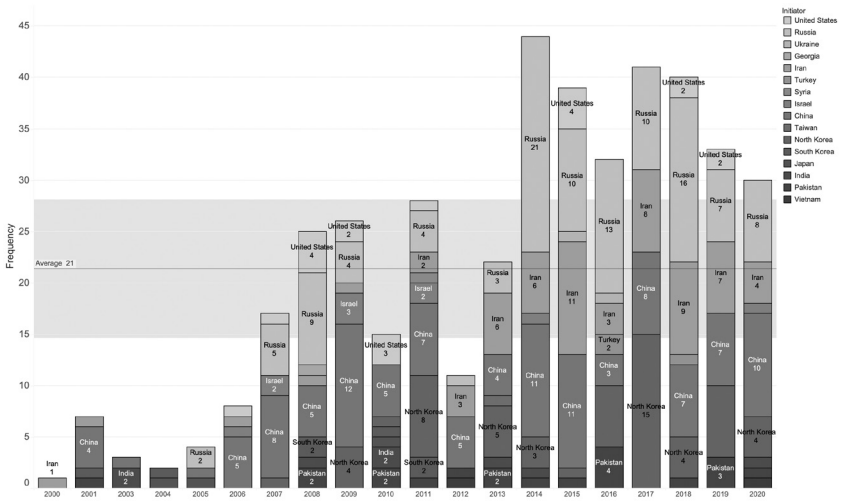

Figure 7. Critical Infrastructure Over Time DCID 2.0

** Note: The researchers found no incidents in the year 2002 and is therefore absent from this time-series graph.*

To provide a time series analysis on CI, Figure 7 visualizes CI incidents by initiator from 2000 to 2020. Across the years, there is an average of about 21 incidents per year among all initiators. In the first decade (2000-2010), China targeted the highest number of CI at

41 out of 108 incidents, or a share of about 38 percent for the entire decade. However, like the trends in Figure 6, Russia increased its targeting of CI from 2011 to 2020 to 93 incidents representing about 29 percent of the 321 incidents recorded for that decade. China followed closely with 73 incidents (23 percent), next Iran with 59 incidents (18 percent), and finally North Korea with 53 incidents (17 percent).

In the first decade of the 21st century, Critical Infrastructure experienced 108 incidents compared to 320 in the second decade. This is a 196 percent increase and further highlights the investments made by China, Russia, Iran, North Korea, and the U.S. in their capabilities to initiate cyber operations. Further, the data indicate that nefarious actors actively target critical infrastructure, and governments around the globe need to exercise resilience and continue to improve their defenses.

Future Directions and Conclusion

This new dataset provides the basics for research into the efficacy of the strategies of persistent engagement, deterrence, and restraint. Future work includes four key areas of data collection. Combining this data with political speeches and signaling demonstrations, military exercises and conventional military incidents, and other forms of engagement and conflict short of war, as competition and engagement, deterrence, and escalation occur in multiple domains.

It is also important to note that the cyber incident data can link to other existing data, including Department of Justice and FBI indictments and sanctions reported publicly through the Congressional Research Service,[53] Department of Treasury,[54] and the Global Sanctions Data Base[55] to produce the cross-database linkages of possible cause and effect. In the past, some have linked this data to event datasets to expand analysis to different domains.[56] Additional data fields also can provide a richer dataset for analysis, such as the Advanced Persistent Threat (APT) attributed to each attack and the corporate targets of the attack.

The next update of DCID will include 2021-2022, pivotal years as they include the Russian invasion of Ukraine, a key point for cybersecurity scholars, and the rise of ransomware attacks during the Pandemic. The project team will continue the legacy of data collection on cybersecurity, hopefully expanding all hybrid warfare actions in the near future.

DCID 2.0 provides a way for cybersecurity research to undertake quantitative data of cyber incidents for the academic and research communities. These data allow for a social science standard for changes over time periods, trend analysis, and hypothesis testing, where data that lacked ratio qualities did not exist. Although the data are an incomplete record because of the unique aspects of cyber, they do capture a sample useful for the study of deterrence, persistent engagement, and escalation. With these data, we hope qualitative claims can be backed by empirical support rather than speculation. This process,

sometimes called normal science, will allow us to learn the correlations and hypothesize the causation of changes in cyber incident trends. Our hope is the data will converge and clarify arguments with statistical facts, rather than continuing to allow them to diverge into conjecture or gross assumptions that jump centuries.

DISCLAIMER

The views expressed in this work are those of the author(s) and do not reflect the official policy or position of the United States Military Academy, the Department of the Army, or the Department of Defense.

NOTES

1. This research would not be possible without our Virtual Student Federal Service data collection team which includes Katrina Villacisneros, Brian Wheat, Ingrid Fontes, Kim Chamberlain, Bridget Flynn, Thomas Deen, Catherine Smith, Nicholas D'Onofrio, Alexandra Restrepo, Siddharth Suman, Sarah Chen, Zachery Gagnon and Jason Niro. Furthermore, this project includes members of the scholarly community who actively work for the U.S. government and the Department of Defense. Our ability to cover and dissect U.S. directed events is limited. We do not attribute the 20 U.S.-responsible cyber incidents to the U.S. government, and included them as strictly the opinions of private sector actors and organizations. As a resource to the wider cybersecurity community, we provide the tools for others to cover global cyber events and build on this work in a comprehensive manner.
2. Branch, Jordan, "What's in a Name? Metaphors and Cybersecurity." *International Organization* 75, no. 1 (2021): 39-70.
3. Senator Angus King and Representative Mike Gallagher, United States of America Cyberspace Solarium Commission, March 2020, 1, https://www.solarium.gov/report.
4. UKGCHQ, The Cyber Threat, 2016. Accessed February 9, 2022, https://www.gchq.gov.uk/information/cyber-threat.
5. Senator Angus King and Representative Mike Gallagher, United States of America Cyberspace Solarium Commission, March 2020, 1, https://www.solarium.gov/report. Pg. 78, 4.3
6. Valeriano, Brandon and Ryan C. Maness, "The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011" *Journal of Peace Research* 51, no. 3 (2014): 347-360.
7. Maness, Ryan C., Brandon Valeriano, and Benjami.n Jensen, "The Dyadic Cyber Incident and Dispute Data, Version 1.5," (June 2019) https://drryanmaness.wixsite.com/cyberconflict/cyber-conflict-dataset.
8. Poznansky, Michael, and Evan Perkoski, "Rethinking secrecy in cyberspace: The politics of voluntary attribution." *Journal of Global Security Studies* 3, no. 4 (2018): 402-416.
9. See CSIS: Significant Cyber Incidents, 2022, https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents, and CFR: Cyber Operations Tracker, 2022, https://www.cfr.org/cyber-operations/. In addition, CSIS mixes criminal attacks with government-focused and-funded attacks. CFR limits its data to state-sponsored actors, for both accuracy and purpose of its dataset.
10. Carnegie Endowment for International Peace, "Timeline of Cyber Incidents Involving Financial Institutions," 2022, https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline.
11. CSIS' criteria are "cyber-attacks on government agencies, defense and high-tech companies, or economic crimes with losses of more than a million dollars." CFR's criteria are "a database of the publicly known state-sponsored incidents that have occurred since 2005." Moreover, "The tracker only contains data in which the perpetrator, also known as the threat actor, is suspected to be affiliated with a nation-state."
12. See the Correlates of War Project at https://correlatesofwar.org. Singer, J. David. "Correlates of War." *New York* (1979). [Vol. 1 is 1979 and Vol 2. 1980]
13. Valeriano, Brandon, and Ryan C. Maness, "How we stopped worrying about cyber doom and started collecting data." Politics and Governance 6, no. 2 (2018): 49-60.
14. Klein, James P., Gary Goertz, and Paul F. Diehl. "The new rivalry dataset: Procedures and patterns." *Journal of Peace Research* 43, no. 3 (2006): 331-348, and Thompson, William R., "Identifying rivals and rivalries in world politics." *International Studies Quarterly* 45, no. 4 (2001): 557-586.
15. Carlson, Brian, "Top Cybersecurity Statistics, Trends, and Facts," CSO, October 7, 2021, https://www.csoonline.com/article/3634869/top-cybersecurity-statistics-trends-and-facts.html.
16. See Lawfare, Topics: Cybersecurity at https://www.lawfareblog.com/topic/cybersecurity and Texas National Security Review, Topics Cyber at: https://tnsr.org/search/cyber/.
17. For more details about the MITRE ATT&CK methodology, see https://attack.mitre.org/.
18. Palmer, Glenn, Roseanne W. McManus, Vito D'Orazio, Michael R. Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, and Michael J. Soules, "The MID5 Dataset, 2011-2014: Procedures, coding rules, and description." *Conflict Management and Peace Science* (2021): https://doi.org/10.1177/0738894221995743.
19. DCID 1.5, https://drryanmaness.wixsite.com/cyberconflict/cyber-conflict-dataset.
20. National Institute of Standards and Technology, "Critical Infrastructure Resources," December 8, 2021, https://www.nist.gov/cyberframework/critical-infrastructure-resources.

NOTES

21. Palmer, Glenn, Vito D'Orazio, Michael R. Kenwick, and Roseanne W. McManus, "Updating the militarized interstate dispute data: A response to Gibler, Miller, and Little." *International Studies Quarterly* 64, no. 2 (2020): 469-475. Brecher, Michael, Jonathan Wilkenfeld, Kyle Beardsley, Patrick James, and David Quinn (2021). *International Crisis Behavior Data Codebook*, Version 14. http://sites.duke.edu/icbdata/data-collections/.
22. Montgomery, M., B. Jensen, E. D. Borghard, J. Costello, V. Cornfeld, C. Simpson, and B. Valeriano, Cyberspace Solarium Commission Report. (2020), Washington, DC. https://www.solarium.gov/report, Valeriano, Brandon and Benjamin Jensen "Building a National Cyber Strategy: The Process and Implications of the Cyberspace Solarium Commission" in T. Jančárková, L. Lindström, G. Visky, P. Zotz (eds.) Going Viral (Report: 2021, 13th International Conference on Cyber Conflict 2021, NATO CCDCOE Publications, Tallinn, Estonia).
23. Uchill, Joe, "Breach Reporting Requirement Sputters as House Passes NDAA," December 8, 2021, https://www.scmagazine.com/analysis/breach/breach-reporting-requirement-sputters-as-house-passes-ndaa.
24. Borghard, Erica D., and Shawn W. Lonergan, "The logic of coercion in cyberspace." *Security Studies* 26, no. 3 (2017): 452-481. Borghard, Erica D., and Shawn W. Lonergan, "Cyber operations as imperfect tools of escalation." *Strategic Studies Quarterly* 13, no. 3 (2019): 122-145. Kreps, Sarah, and Jacquelyn Schneider, "Escalation firebreaks in the cyber, conventional, and nuclear domains: moving beyond effects-based logics." *Journal of Cybersecurity* 5, no. 1 (2019): tyz007. Healey, Jason, "The implications of persistent (and permanent) engagement in cyberspace." *Journal of Cybersecurity* 5, no. 1 (2019): tyz008, Brantly, Aaron F., ed. *The Cyber Deterrence Problem*. Rowman & Littlefield Publishers, 2020.
25. For more information, please see https://vsfs.state.gov/.
26. Valeriano, Brandon, Benjamin M. Jensen, and Ryan C. Maness, *Cyber strategy: The evolving character of power and coercion*. Oxford University Press, 2018.
27. Healey, Jason, "Beyond Attribution: Seeking National Responsibility in Cyberspace," Atlantic Council, February 22, 2012, https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/beyond-attribution-seeking-national-responsibility-in-cyberspace/.
28. Maness, Ryan C., Brandon Valeriano, and Benjamin Jensen, "The Dyadic Cyber Incident and Dispute Dataset (DCID) Version 1.5", Codebook, 2019: https://drryanmaness.wixsite.com/cyberconflict/cyber-conflict-dataset.
29. National Institute of Standards and Technology, "Critical Infrastructure Resources," December 8, 2021, https://www.nist.gov/cyberframework/critical-infrastructure-resources.
30. Whyte, Christopher, Brandon Valeriano, Benjamin Jensen, and Ryan Maness. "Rethinking the data wheel: Automating open-access, public data on cyber conflict." In *10th International Conference on Cyber Conflict (CyCon)*, 9-30. IEEE, 2018.
31. Cornell Law School, "44 U.S. Code § 3552- Definitions," https://www.law.cornell.edu/uscode/text/44/3552.
32. CFR uses the term "operation," while CSIS uses the term "incident."
33. Cornell Law School, "44 U.S. Code § 3552 - Definitions," https://www.law.cornell.edu/uscode/text/44/3552.
34. Wolff, Josephine, "Why We Need to Be Much More Careful About How We Use the Word *Cyberattack*," *Slate*, March 30, 2017, accessed February 8,2022, https://slate.com/technology/2017/03/we-should-be-careful-when-we-use-the-word-cyberattack.html
35. Trapdoors or Trojans are unauthorized software added to a program to allow entry into a victim's network or software program to permit future access to a site once it has been initially attacked. The purpose of trapdoors is to steal sensitive information from secured sites.
36. (1) Logic bombs, or wiper malware, are programs that cause a system or network to shut down and/or erase all data within that system or network. (2) Viruses are programs which attach themselves to existing programs in a network and replicate themselves with the intention of corrupting or modifying files. (3) Worms are essentially the same as viruses, except they do not need to attach themselves to existing programs. (4) Keystroke logging is the process of tracking the keys being used on a computer so that the input can be replicated in order for a hacker to infiltrate secure parts of a network.
37. Foote, Colin, Ryan Maness, Benjamin Jensen, and Brandon Valeriano, "Cyber conflict at the intersection of information operations: Cyber-enabled information operations, 2000-2016," in *Information Warfare in the Age of Cyber Conflict* (Routledge, 2020), 54-69.
38. Valeriano et al., 2018, *Cyber Strategy*.
39. For detailed information on the DCID severity scale and events coded 7-10, see Maness, Ryan C., Brandon Valeriano, Kathryn Hedgecock, Jose M. Macias, and Benjamin Jensen, "The Dyadic Cyber Incident and Campaign Dataset, version 2.1, *Codebook*, 2022, https://drryanmaness.wixsite.com/cyberconflict/cyber-conflict-dataset.

NOTES

40. Rid, Thomas, and Ben Buchanan, "Attributing cyber-attacks." *Journal of Strategic Studies* 38, no. 1-2 (2015): 4-37.
41. See 16 critical infrastructure sectors at https://www.cisa.gov/critical-infrastructure-sectors.
42. For a complete breakdown of CI please see Maness et al., 2022.
43. In multiple locations the authors previously have discussed the limitations of coding cyber incidents in the public knowledge and have pointed the reasons why this is possible despite assumptions to the contrary. Due mainly to the importance of the topic, the awareness of the news media, and the rush to provide information to trumpet the ability of cybersecurity incident response firms, much information is available on cyber operations despite most being covert operations.
44. Valeriano et al., 2018, *Cyber Strategy*.
45. Valeriano et al., 2018, *Cyber Strategy*.
46. Department of Defense, DoD Cyber Strategy 2018, https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBER_STRATEGY_SUMMARY_FINAL.PDF.
47. Here we summarize severity level six, "Single/multiple critical network infiltration and widespread destruction" from our codebook as strategic sabotage.
48. Valeriano, Brandon, and Ryan C. Maness. *Cyber war versus cyber realities: Cyber conflict in the international system.* Oxford University Press, USA, 2015.
49. DiMaggio, John, "A History of REvil," *Analyst 1,* 2022, https://analyst1.com/file-assets/History-of-REvil.pdf
50. McAfee, "What is Petya and NotPetya Ransomware?" 2017, https://www.mcafee.com/enterprise/en-us/security-awareness/ransomware/petya.html.
51. BBC News, "North Korea hackers stole $400m of cryptocurrency in 2021, report says," bbc.com, January 14, 2022, available at https://www.bbc.com/news/business-59990477#:~:text=Separately%2C%20in%20February%20last%20year,the%20Department%20of%20Justice%20said.
52. Berr, Jonathan, "'WannaCry' ransomware attack losses could reach $4 billion," *CBS News,* May 16, 2017, available at https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/.
53. See, for example, Katzman, Kenneth, "Iran Sanctions (RS20871)," Congressional Research Service, 2022, https://crsreports.congress.gov/product/details?prodcode=RS20871.
54. U.S. Department of Treasury, "Consolidated Sanctions List (Non-SDN lists)," December 16, 2021, https://home.treasury.gov/policy-issues/financial-sanctions/consolidated-sanctions-list-non-sdn-lists.
55. Felbermayr, Gabriel, Aleksandra Kirilakha, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov, "The global sanctions data base." *European Economic Review* 129 (2020): 103561.
56. Maness, Ryan C., and Brandon Valeriano, "The impact of cyber conflict on international interactions." *Armed Forces & Society 42*, no. 2 (2016): 301-323, Valeriano et al., 2018, *Cyber Strategy*.