2019 11th International Conference on Cyber Conflict:
Silent Battle
T. Minárik, S. Alatalu, S. Biondi,
M. Signoretti, I. Tolga, G. Visky (Eds.)
2019 © NATO CCD COE Publications, Tallinn

Rough-and-Ready: A Policy Framework to Determine if Cyber Deterrence is Working or Failing

Jason Healey
Senior Research Scholar
Columbia University, SIPA
New York, NY USA
jh3639@columbia.edu

Neil Jenkins
Chief Analytic Officer
Cyber Threat Alliance
Arlington, VA USA
neiljenkins@cyberthreatalliance.org

Abstract: This paper addresses the recent shift in the United States' policy that emphasizes forward defense and deterrence and to "intercept and halt" adversary cyber operations. Supporters believe these actions should significantly reduce attacks against the United States, while critics worry that they may incite more adversary activity. As there is no standard methodology to measure which is the case, this paper introduces a transparent framework to better assess whether the new U.S. policy and actions are suppressing or encouraging attacks.[1]

Determining correlation and causation will be difficult due to the hidden nature of cyber attacks, the veiled motivations of differing actors, and other factors. However even if causation may never be clear, changes in the direction and magnitude of cyber attacks can be suggestive of the success or failure of these new policies, especially as their proponents suggest they should be especially effective. Rough-and-ready metrics can be helpful to assess the impacts of policymaking, can lay the groundwork for more comprehensive measurements, and may also provide insight into academic theories of persistent engagement and deterrence.

Keywords: *cyber deterrence, metrics, cyber conflict, cyber operations, threat intelligence, cyber policy, persistent engagement*

1 This work was funded in part by the Office of Naval Research under the OSD Minerva program: Grant number N00014-17-1-2423.

1. INTRODUCTION

The United States has significantly shifted its policy regarding the Department of Defense (DoD)'s role in cyberspace to emphasize "persistent presence," to remain in "in foreign cyberspace to counter threats as they emerge" and to "intercept and halt cyber threats."[2] The belief is that, over time, these actions will cause nation state adversaries (particularly Russia, China, Iran, and North Korea) to become less effective; they will be forced to expend more resources on defense and will choose not to engage the United States.

Beyond such active engagement with adversaries, the new policy also seeks to impose costs through deterrence, both outside and inside cyberspace. The measures outside cyberspace include actions like sanctions and indictments, while those inside include gaining access to systems that adversaries value, to hold them at risk with offensive cyber operations. In the words of John Bolton, the National Security Advisor, the White House has "authorized offensive cyber operations […] not because we want more offensive operations in cyberspace, but precisely to create the structures of deterrence that will demonstrate to adversaries that the cost of their engaging in operations against us is higher than they want to bear."[3]

Supporters believe these are long-awaited steps which will significantly reduce transgressions against the United States. Critics believe such counteroffensive activities may only inflame nation state adversaries, who could see them not as a mild corrective but as a fresh insult which demands a response. There is currently no standard methodology to measure whether the new U.S. policy and actions are suppressing or encouraging attacks. While it would be routine for a military command like U.S. Cyber Command to have measures of effectiveness for specific military operations, this is not necessarily true for assessments of the policy outcomes. To this end, Representative James Langevin (D, RI-2) is pushing for such measures: "Much like the traditional battlefield, we must measure the impact of our operations to assess our warfighting effectiveness towards the larger objectives and ensure our strategic vision reflects the realities of engagement in cyberspace."

2 Nakasone, Paul M. 2019. "An Interview with Paul M. Nakasone," Joint Forces Quarterly. https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-92/jfq-92.pdf.

3 Bolton, John. 2018. "Transcript: White House Press Briefing on National Cyber Strategy - Sept. 20, 2018." Washington DC (September 8). Available at https://news.grabien.com/making-transcript-white-house-press-briefing-national-cyber-strateg.

4 Langevin, James R. 2019. "Opening Statement: FY 2020 Budget Request for Military Operations in Cyberspace." March 13. https://armedservices.house.gov/_cache/files/d/5/d5f94725-3373-40ef-803c-1f0ff8f106a8/577D710BF48F37825B2656EE1AF6891A.opening-statement---ietc-chairman-langevin-3-13-2019.pdf.

This paper is intended to help bridge this gap and has four related goals:

1. *Stimulate the conversation*. Despite significant commentary and research on the new policy, there has been little discussion on how to assess if it is working as expected or not.
2. *Propose basic metrics* which might suggest if the new policy is working as expected to dissuade attacks or is actually encouraging them. Even simple metrics might make some causal explanations more or less likely, even though determining strong correlation (much less causation) may be distant goals.
3. *Introduce a basic framework* of terms and concepts. Security threat analysts, policymakers, and researchers need an analytical structure to make it easier to weigh evidence and make conclusions.
4. *Encourage more complex, data-driven approaches* from those who may be able to determine causation or correlation, such as the U.S. Intelligence Community and the commercial cybersecurity threat intelligence community.

It is obviously hard to prove whether any kind of policy to influence adversaries is working or not. It is still not definitively settled if the lack of Cold War nuclear attacks between the United States and the Soviet Union was the result of deterrence or a lucky coincidence. It has been three years since the U.S. and Chinese presidents agreed to limit cyber espionage for commercial benefit, and the cyber-threat and policy communities continue to debate if the Chinese did in fact reduce such espionage and whether any such changes were meaningful or due to the agreement (or other U.S. actions such as indictments).[5] For both nuclear and cyber attacks, it is fortunately easier to measure failure than success. Successful policies may succeed quietly but fail explosively. A skyrocketing increase in Chinese espionage operations after the Obama-Xi agreement would have been a far clearer signal than the apparent decrease.

The metrics in this paper have obvious shortcomings: they cannot prove causality and cannot usually be based on specific U.S. actions, which will likely be classified – such as threats expressed privately to adversaries or counter-offensive disruptions by U.S. Cyber Command. Usually, only the overall policy and pace of adversary attacks will be known, at least from public sources. Attribution and adversaries' decision calculus in many cases cannot be understood quickly. These shortcomings can all be minimized with the right framework and by comparison with additional data sources to address key questions.

Rough-and-ready metrics, including determining the direction and magnitude of any changes over time, are needed to assess the impacts of cyber policymaking. The U.S. military is already conducting these operations, so policymakers need good enough

5 Segal, Adam. 2016. "The U.S.-China Cyber Espionage Deal One Year Later." Council on Foreign Relations, September 28. https://www.cfr.org/blog/us-china-cyber-espionage-deal-one-year-later.

metrics now to assess the overall effort.[6] With only marginal changes in terminology, this framework can also be useful for efforts to advance "digital peace," such as those by Microsoft and France.[7]

This paper proceeds by first introducing the new U.S. government policies and examining issues of measurement. Then, it discusses several frameworks, starting from simple, illustrative examples, to more fuller descriptions of categories of transgressions. It addresses shortcomings of the framework before a short conclusion and recommendation for future work.

2. THE DOD POLICIES

The new DoD strategy is based on "persistent presence", in part to "intercept and halt" adversary operations – imposing costs on their *current* operations – as well as outright deterrence so that they will choose not to undertake *future* operations, as they fear the costs imposed by the United States will be "higher than they want to bear."[8]

As an example of what this might mean in practical terms, if Iranian cyber operators were gathering resources to conduct further disruptive campaigns against the United States financial sector (as they did in 2011-2012), U.S. Cyber Command could seek to disrupt their efforts in foreign cyberspace, up to and including counter-offensive cyber operations.[9] In the short term, this would impose "tactical friction", dissuading the Iranians as they have to defend themselves and expend resources to rebuild the disrupted capabilities and infrastructure. These operations for persistent presence and forward defense would be heightened with actions specifically aimed at cyber deterrence, such as U.S. Cyber Command holding Iranian critical infrastructure at risk of a counter-attack with offensive cyber operations.

While such actions for persistent presences are an innovation in cyber conflict, applying concepts of deterrence is not. A recent Defense Science Board task force characterized cyber deterrence as actions "affecting the calculations of an adversary … to convince adversaries not to conduct cyber attacks or costly cyber intrusions."[10] Deterrence in cyberspace is a complex web of deterrence by denial (actions that reduce

6 Barnes, Julian E. 2018. "U.S. Begins First Cyberoperation Against Russia Aimed at Protecting Elections." The New York Times, October 23. https://www.nytimes.com/2018/10/23/us/politics/russian-hacking-usa-cyber-command.html.

7 See Microsoft, Digital Peace, https://digitalpeace.microsoft.com/; and Paris Peace Forum, held November 2018, https://parispeaceforum.org/.

8 Nakasone. 2019. Bolton. 2018.

9 Perlroth, Nicole, and Quentin Hardy. 2013. "Bank Hacking Was the Work of Iranians, Officials Say." The New York Times, January 8. https://www.nytimes.com/2013/01/09/technology/online-banking-attacks-were-work-of-iran-us-officials-say.html.

10 Defense Science Board, Department of Defense. 2017. "Task Force on Cyber Deterrence." Defense Science Board, 3, 4. https://www.acq.osd.mil/dsb/reports/2010s/DSB-cyberDeterrenceReport_02-28-17_Final.pdf.

the effectiveness of attacks, most notably by improving cyber defenses, network protection and security, and resilience) and deterrence by cost imposition (actions that increase the costs of the adversary when attacking, such as public attribution and shaming, diplomatic actions, economic sanctions, and the risk of a cyber or physical counterattack).[11]

The U.S. Government has used these components of deterrence over the last several years, with various levels of success. One publicly available report notes that network defense alone: "will not be sufficient to deter determined and sophisticated state-sponsored adversaries" and "the United States will also undertake a new effort to increase deterrence of state actors through cost imposition and other measures".[12] The new policy accordingly joins typical defense actions (like information sharing), with specific deterrent actions and operations for persistent presence, not meant to deter future attacks but to disrupt those underway or expected.

The assessment of this newly forceful DoD policy has broadly split into two camps, which we dub "hawks" and "owls." The hawks accept that a more forceful U.S. response with offensive cyber operations will work the way Bolton predicts, imposing "negative feedback" leading to a reduction in transgressions by adversaries. The owls are more cautious, worried that offensive cyberattacks – even if justified – may instead create "positive feedback," inciting more attacks in return.

There is sparse evidence supporting either position and the debate on whether the new policy will garner negative or positive feedback will not be settled through discussion and opinion, no matter how many op-eds are written. Rather, analyzing the success or failure of the new policy requires an evidence-based approach with a repeatable and transparent framework.

3. MEASURING CYBER DETERRENCE AND PERSISTENT ENGAGEMENT … INDIRECTLY

There have been few, if any, significant efforts to comprehensively measure these effects of persistent engagement or cyber deterrence on adversary behavior.

Scholars and practitioners have perhaps been dissuaded because the discussion for the past decade has been focused on cyber deterrence and it is "virtually impossible

11 Note that cyber deterrence in this context only applies to a limited subset of adversaries: those tied to nation states, especially Russia, China, Iran, and North Korea (and any proxy or irregular groups supported by states). It does not apply to non-state groups like criminals or hacktivists.

12 Department of State. 2018. "Recommendations to the President on Deterring Adversaries and Better Protecting the American People from Cyber Threats." May 31. https://www.state.gov/documents/organization/282253.pdf.

to know if deterrence is working" in cyberspace.[13] When the discussion is framed as deterrence, Henry Kissinger stated what is still a common position: "Since deterrence can only be tested negatively, by events that do not take place, and since it is never possible to demonstrate why something has not occurred, it became especially difficult to assess whether the existing policy was the best possible policy or just a barely effective one."[14] Yet, this is mostly true only in measuring the success of deterrence. Its failure would have been obvious soon after detonation of the first atomic warheads in an enemy's heartland; but because nuclear war never happened, this was not a practicable distinction.

The upside for cyber conflict is that – unlike with nuclear weapons – engagements and campaigns are constantly happening in cyberspace: what DoD is now calling "persistent engagement". These operations are tracked over time – though not to determine the success of different policies, such as persistent engagement and deterrence. The downside is that much of the activity of engagement and campaigns is hidden, cause and effect blend and overlap, and the identity of the adversary is obscured.[15]

Past cyber incidents show a range of "knowability" of the impact of adversary actions. On the more knowable end of that range, in responding to election interference in 2016, the administration of President Barack Obama took response actions off the table out of concern that Russia would escalate "against America's critical infrastructure—and possibly shut down the electrical grid" or engage in "hacking into Election Day vote tabulations."[16] This is knowable because principals involved in the Situation Room themselves confirmed the impact of Russian capabilities.

The second case, the 2015 agreement by President Barack Obama and President Xi Jinping of China, is not as clear; but, as will be discussed below, the debate can be addressed with data and an analytical framework. It is accordingly far more tractable than determining any effects of President Obama's warning to President Vladimir Putin over election interference. After the warning, in Hangzhou, China in September 2016, the U.S. government detected "no further evidence of Russia cyber-intrusions

13 Sulmeyer, Michael. 2018. "How the U.S. can Play Cyber-Offense." Foreign Affairs, March 22. https://www.foreignaffairs.com/articles/world/2018-03-22/how-us-can-play-cyber-offense.

14 Kissinger, Henry. 1994. Diplomacy. Simon and Schuster. p608.

15 In the Cold War, academics and policymakers generally knew far more about U.S. operations and capabilities than those of the Soviets. In cyber conflict, the reverse is true: the DoD and Intelligence Communities publicly discuss adversary operations against the United States, while highly classifying their own operations against the same adversaries, masking critical issues such as determining of cause and effect.

16 Healey, Jason. 2018. "Not the Cyber Deterrence the United States Wants." Council on Foreign Relations, June 11. https://www.cfr.org/blog/not-cyber-deterrence-united-states-wants. Subsequently confirmed in conversation with Gen. Clapper, former Director of National Intelligence; and Peter Clement, Columbia University, 21 February 2019.

into state election systems."[17] Such systems are typically not as tightly monitored as corporate networks, so there may be little data from which to draw conclusions. Barring exquisite intelligence or confirmation by the Kremlin, determining whether the warning caused any decrease is highly problematic.

Despite these drawbacks, this newly muscular U.S. strategy, like all policies, needs to be measured as best as possible to determine its effectiveness. Although they may not be definitive, rough-and-ready measurements of the scope and number of cyber incidents can suggest the impact of persistent engagement and deterrence. There has been some work in this space, especially by Brandon Valeriano and Ryan Maness, but it has been focused on academic questions of deterrence and not yet on policy effectiveness.[18]

The simplest metric framework is to describe different levels and types of cyber transgressions and simply tot them up. The next paragraphs describe such examples; these can be used as mere illustrations of the concept, especially to highlight that big data sets are not required (and may just overcomplicate the analysis, hiding more obvious signals), but can also be reasonable frameworks in their own right. Each metric should be tied directly to the goals of the policymakers.

A. Three Basic Frameworks

The Federal Government uses a standard five-tier severity score, the Cyber Incident Severity Schema, to assess the gravity of incidents.[19] The Schema rates cyber incidents according to observed effect, impact, affected sectors, and attribution (if known). Users can make qualitative judgments on these categories or attempt to score and weight them. If the general policy goal is to impose costs on adversaries and reduce the number and scope of significant incidents, this could be operationalized by tracking the number of attacks rated level 3 and above (those that are "significant cyber incidents" and likely to result in impacts to "public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence"). The metric is then a simple algorithm along the lines of, "*if* cyber_level = {3, 4, 5} then count = count +1," tracked over time.[20] This metric might also be tied to the higher threshold of "use of force" to fit more neatly with the stated goal of the DoD

17 Isikoff, Michael, and David Corn. 2018. "'Stand Down': How the Obama Team Blew the Response to Russian Meddling." Huffington Post, March 9. https://www.huffingtonpost.com/entry/stand-down-how-the-obama-team-blew-the-response-to-russian-meddling_us_5aa29a97e4b086698a9d1112.

18 Valeriano, Brandon, and Ryan C Maness. 2015. Cyber War versus Cyber Realities. Cyber Conflict in the International System. Oxford University Press.

19 U.S.-CERT, Department of Homeland Security. n.d. "NCCIC Cyber Incident Scoring System." https://www.us-cert.gov/NCCIC-Cyber-Incident-Scoring-System. Brandon Valeriano has pointed out that in his experience, a 10-point scale would allow finer grained analyses; but for now, DHS uses only five points.

20 The authors intend to use the scheme to assess a number of past examples, possibly including Target, Iranian DDoS of the finance sector, Shamoon, Sony, OPM, Ukrainian power outage/Black Energy, Russian election interference, and Cloud Hopper.

strategy: to prioritize "deterring malicious cyber activities that constitute a use of force against the United States, our allies, or our partners."[21]

As attribution becomes clearer for each incident, this metric becomes more useful, as there can be a separate count for each of the United States' major adversaries in cyberspace, especially China, Russia, Iran, and North Korea. As the U.S. Government already assesses this score for all incidents to which they respond, there is no additional cost to tracking the trends over time to determine if the new policy has measurable impact. It is most useful for tracking discrete events, such as denial of service attacks or malware outbreaks (like NotPetya) and individual espionage incidents (OPM) than for less easily counted espionage campaigns (Cloud Hopper) or implanting malicious software for future use (Havex/Black Energy).[22]

This may be sound overly simple, but based on our knowledge and interviews, such tracking and measurement is less routine than may be imagined. Even marginal gains can be immediately useful to policymakers.

Bolton explained that the 2015 intrusion by China into the Office of Personnel Management was just "the kind of threat to privacy from hostile foreign actors that we're determined to deter".[23] This policy goal can be operationalized to a rough-and-ready metric framework by developing three elements:

1. A general description of an "OPM-type" incident, such as by scope, duration, intensity or against international laws or norms, U.S. red lines, or explicit agreement between states.
2. A measured baseline of such incidents. These can largely be drawn from headlines as they are both relatively few in number and quickly become public knowledge.
3. Tracking new developments to see whether the number of "OPM-type" incidents increases or decreases after the new policy comes into force.

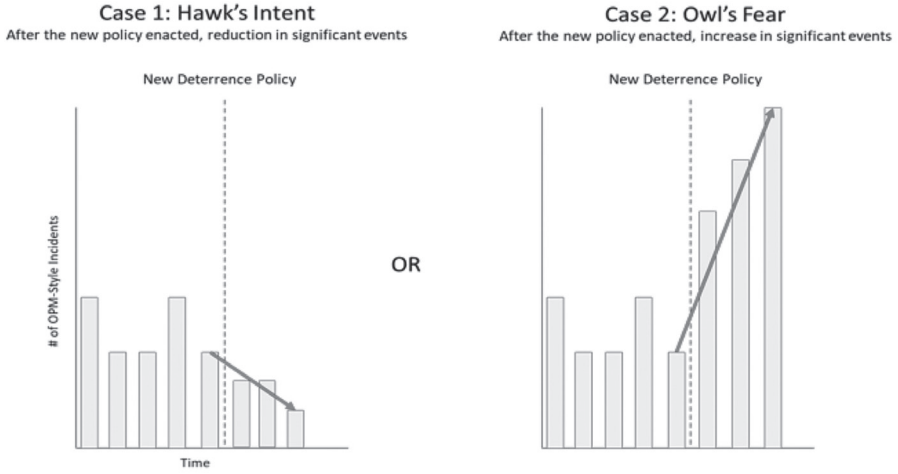
Although this simple metric could not meaningfully "prove" whether or not Bolton's threatened deterrence worked, if there was a decrease in such incidents (see case 1), then the evidence might indeed support Bolton's policy. But if there is a sharp increase in OPM-style incidents (see case 2), this suggests that the policy might be counter-productive. Further analysis is needed to check the competing hypotheses. The increase in case 2 is particularly significant, as the hawks suggest that the new U.S. policies

21 Defense Science Board. 2017.

22 NotPetya was the devastating 2017 ransomware outbreak caused by Russia, while OPM refers to the espionage incidents involving the U.S. Office of Personnel Management in 2015. Cloud Hopper was a large-scale Chinese espionage operation against managed service providers. Havex/Black Energy were malware implanted widely in energy grids but not, except for the notable exception of Ukraine, actually used to cause disruption.

23 Bolton. 2018.

and actions should have a substantial impact on adversary operations. Therefore, any movement of the trend in the opposite direction has far more significance. Failure is louder than success.



The above two frameworks rely on data more than context. A more applied framework centers on measuring the effects of U.S. policies and actions meant to deter or dissuade a specific adversary from a specific kind of transgression. This more detailed metric is particularly useful for campaigns and implants, as it tracks, rather than individual incidents, volumes of activity over time from a specific adversary. The best example is the FireEye assessment that there had been a "notable decline in China-based groups' overall intrusion activity against entities in the U.S. and 25 other countries", with an especially sharp decline after the Obama-Xi agreement in 2015.[24] That company's cyber threat intelligence analysts measured the number of "active networks compromised" by 72 suspected China-based groups (see Chart x below).

The 90+ percent decrease, according to FireEye, was due to "ongoing political and military reforms in China, widespread exposure of Chinese cyber activity, and unprecedented action by the U.S. government." It is noteworthy that two of these three reasons are related to U.S. counters: the public naming-and-shaming of "widespread exposure"; and "unprecedented" indictments and the threat of sanctions.

24 FireEye iSight Intelligence. 2016. "Redline Drawn: China Recalculated Its Use of Cyber Espionage." FireEye. https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/rpt-china-espionage.pdf.

There is still debate about three key issues:

1. Was there was an actual decrease in Chinese espionage operations for commercial purposes? Perhaps the number of incidents held steady but the bulk were not detected, due to improved Chinese stealth. This is generally a question for cyber threat analysts.
2. How much of any Chinese response was the result of the U.S. policy? Perhaps the Chinese primarily acted for their own reasons, in response to domestic Chinese pressures, and U.S. policies had little additional impact. This is a question best answered by China experts.
3. Did the decrease matter? Perhaps the few networks still being compromised were those most critical to national security, so the overall impact was not meaningfully diminished. This is a question best answered by the policymakers themselves.

It has been over three years since the Obama-Xi agreement, yet there has been little if any structured work that has pulled out and analyzed these separate strands, rather than addressing whichever one supports the authors' preconceived ideas about China or the efficacy of agreements.

The evidence to answer these three questions, if not definitive, is certainly suggestive. We know many of our colleagues will disagree, some vehemently, with these assessments. This only reinforces the key point: that estimative conclusions must be systematically addressed in a transparent framework and tied to policymakers' goals. Isolating each element allows more transparency and repeatability, so that different analysts with different sources of information can develop individual and collective assessments of whether U.S. policies and actions are working or not.

Assistant Attorney General John Carlin confirmed FireEye's assessment of the direction and magnitude of Chinese activity: "Consistent with their agreement, they largely ceased state-sponsored hacking that targeted a private US company for the direct economic benefit of a Chinese competitor."[25] Even as late as November 2018, Rob Joyce, the former White House cybersecurity coordinator and NSA executive, with access to unique sources of intelligence, felt that although Chinese activity had returned, it had still dropped "dramatically" since the agreement of three years before.[26] We assess with medium confidence that it is very likely that there was a significant drop in Chinese activity.

On whether U.S. pressure worked, Xi has indeed been centralizing power in the Communist Party and his own person while cracking down on corruption. Either or both drives may have led Xi to clamp down on barely authorized cyber operations for commercial purposes. Even so, the Chinese, Carlin felt, "saw they had a big potential embarrassment brewing", while another Justice official noted "they were highly motivated to do the right thing".[27] According to Michael Daniel, then the White House cyber coordinator, the agreement was due to "steady, sustained pressure through a number of channels, including direct diplomacy, indirect diplomatic activity, public statements, and law enforcement actions" and "the Chinese were also concerned about potential additional actions that the U.S. could have taken, such as economic sanctions".[28] In addition, "President Xi had an upcoming visit to the United States and the Chinese wanted to make cybersecurity a positive topic, rather than a source of tension during the visit". Yet, with high confidence, we analyze it as very likely that U.S. pressure helped push Xi's decision – and subsequent Chinese action – in the preferred direction.

Regarding whether the decrease mattered, much of the original U.S. complaint was not only that the Chinese were stealing secrets for commercial purposes, but that the sheer quantity of such transgressions themselves was destabilizing. This is not an intelligence assessment but a policy judgment; but we believe this was a win for the United States. A reduction in the number of incidents directly relates to a decrease in perceived aggression: fewer incidents affected fewer companies. This is a strong benefit, in line with the U.S. policy goals, even if there was an impact from the remaining incidents. For other kinds of transgressions (see below), the policy goal must be not just fewer incidents, but zero, with any adversary action unacceptable. Espionage for commercial purposes is not usually such a zero tolerance issue.

25 Graff, Garrett M. 2018. "How the US Forced China to Quit Stealing—Using a Chinese Spy." Wired, October 11. https://www.wired.com/story/us-china-cybertheft-su-bin/.

26 Reuters. 2018. "US Accuses China of Violating Bilateral Anti-Hacking Agreement." CNBC, November 8. https://www.cnbc.com/2018/11/09/us-accuses-china-of-violating-bilateral-anti-hacking-agreement.html.

27 Graff. 2018.

28 Daniel, Michael. 2019. Interview with Michael Daniel (January 8).

B. The Model Applied to Other Transgressions

As illustrated in the China example above, measurement frameworks work best when the policy goals are clearly stated. As cyber incidents can take so many forms, this next section will articulate different kinds of transgressions, both to simplify the lexicon for policymakers and to define different categories for measurement. Analysts can, as above, rate the severity of transgressions and assign these to one or more categories. These are samples: there may be a larger set, especially as technology and adversary attacks develop over the decades.

Reckless Incidents: Some cyber transgressions have shown a "lack of regard for the danger or consequences," falling well outside the norms and having global effects.[29] These include attacks that have cascading or systemic effects, which cause significant cyber effects well beyond the intended target or original goal; or attacks that largely only affect their intended target, but that target itself is particularly critical or with a high potential for mistake or miscalculation and possibility of massive damage. As the true intent of the attacker may not be known, this will often be an analytical judgement based on effects and impact.

Coding, in the social sciences, means to apply categories to facilitate analysis. To determine if an incident should be coded as "reckless," how widespread the disruptive effects were (such as local to intended target, regional outside of intended target, or global), or the sensitivity or criticality of the intended target, might be assessed. For example, the NotPetya and WannaCry attacks, from Russia and North Korea respectively, both had globally disruptive impact. The Chinese "Great Cannon" denial of service against Github affected not only that software repository site, but developers globally who depended upon it.[30] All might be coded as "reckless."

Brazen Incidents: As seen in Bolton's response to the OPM espionage incident, some cyber incidents have a scope, duration and intensity that necessitates a significant national security response by the attacked nation: "This must not stand".[31] Some of these brazen attacks may cross a specific threshold, such as causing death or physical destruction or defying international law and norms. But "brazen" is not a legal threshold but a political judgment, as even espionage could be brazen if of an appropriate scope, duration, or intensity. As with "reckless," the intent of adversaries cannot be known and what might seem "brazen" to the defender might seem reasonable (or even just deserts) to others.

Possible coding for brazen transgressions includes the number of deaths; a measure of disruption or destruction (such as economic cost or number of systems "bricked");

29 Oxford. 2009. "Recklessness." In Oxford English Dictionary. Oxford University Press.

30 Marczak, Bill, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. 2015. "China's Great Cannon." The Citizen Lab. April 10. https://citizenlab.ca/2015/04/chinas-great-cannon/.

31 Thanks to Christopher Painter for the recommendation to tie "brazen" to incidents that necessitate a response.

and whether the incident violated a norm previously agreed to by the attacker, a global norm, a national "red line," or none of these. The Chinese intrusion into OPM has been mentioned above as a possible brazen attack, while Russian interference in the 2016 U.S. elections is an even more obvious candidate.[32] The U.S. has conducted its own brazen attacks, most notably the Stuxnet malware attack (conducted with Israel) against Iranian uranium enrichment.[33]

Destabilizing Presence: Some systems are so critical and hazardous that *any* foreign cyber presence is extraordinarily high-risk and potentially destabilizing. For example, gaining access to the command and control systems of a nation's nuclear weapons could precipitate a nuclear war. Access to the control systems of a nuclear power plant or a massive dam could be similarly high-risk, as even a simple key stroke error could cause a disaster. To a lesser degree, gaining access and pre-positioning malware in another nation's electrical grid could be destabilizing because it could lead to a sudden, strategic strike.[34]

Possible coding for this category is far simpler, as it depends on the degree of adversary presence: zero, limited, or widespread. Perhaps the most worrying example is the Black Energy and Havex malware implanted by Russia in U.S. and European electrical systems, including nuclear power plants.[35] A variant was subsequently used to disrupt the Ukrainian power grid, in what was certainly also a brazen incident, highlighting that these categories are not mutually exclusive.[36]

Disproportionate Response: Nations are frequently subjected to intrusions and low-level disruption from other states. Another kind of transgression, related to those above, is when a nation's response is far out of scale to the harm done to it. This is likely a small category, but is included here as the policy response to a disproportionate response should be different from the response to a pure brazen attack.

Possible coding for disproportionate response could include comparing the level of the initial incident (such as number of systems disrupted) with the response. For example, it is possible that the Iranians conducted the large-scale Shamoon attack

32 Office of the Director of National Intelligence. 2017. "Background to 'Assessing Russian Activities and Intentions in Recent US Elections': The Analytic Process and Cyber Incident Attribution. https://www.dni.gov/files/documents/ICA_2017_01.pdf.; Sanger, David E. 2018. The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age. Crown. pxviii.

33 Zetter, Kim. 2014. "An Unprecedented Look at Stuxnet, the World's First Digital Weapon." Wired, November 3. https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/.

34 Clarke, Richard A, and Robert K Knake. 2011. Cyber War: The Next Threat to National Security and What to Do About It. Ecco. p244.

35 F-Secure Labs. 2014. "BlackEnergy & Quedagh: The Convergence of Crimeware and APT Attacks." https://www.f-secure.com/documents/996508/1030745/blackenergy_whitepaper.pdf. Constantine, Lucian. 2014. "New Havex Malware Variants Target Industrial Control System and SCADA Users." PC World, June 24. www.pcworld.com/article/2367240/new-havex-malware-variants-target-industrial-control-system-and-scada-users.html.

36 Jackson Higgins, Kelly. 2016. "Lessons from the Ukraine Electric Grid Hack." Dark Reading, March 18. https://www.darkreading.com/vulnerabilities---threats/lessons-from-the-ukraine-electric-grid-hack/d/d-id/1324743.

against Saudi Aramco and Qatari Rasgas because their own energy infrastructure had been hit by a similar Wiper worm only weeks beforehand.[37] It seems likely to have been a disproportionate response rather than a fresh transgression. Distinguishing tit from tat is an important analytical distinction.

Attacker Infrastructure: This category stands apart, as it does not capture the output metrics of actual transgressions, but the impact of U.S. operations on adversary attack infrastructure – hop points, command and control servers, development and test environments, capabilities and the like. One hope for the new U.S. cyber doctrine is for U.S. operations to have a "strategic effect as the 'tactical friction' the adversary experiences through continuous engagement by the United States compels them to shift their resources (and thinking) toward their own vulnerabilities and defense."[38] This implies, in part, operations against adversary attack infrastructure to impose that friction, which can be directly measured. It would not be surprising if U.S. Cyber Command were using such measures to assess the effectiveness of their actions, but this can also be tracked by commercial cybersecurity companies.

Mandiant, in its groundbreaking report on the APT1 group, noted that the Chinese espionage team had "937 Command and Control (C2) servers hosted on 849 distinct IP addresses in 13 countries," and were "logging into their attack infrastructure from 832 different" Internet addresses.[39] Tracking these same metrics over time allows a rough measure of U.S. operational effectiveness. This could include total infrastructure disrupted, by category and by ratio of the total known infrastructure, and the mean time to rebuild.

C. Addressing the Shortcomings of This Approach

It can be relatively straightforward to use this framework as a rough-and-ready measure of the effects of U.S. policies and actions. This requires the three steps mentioned in the OPM example above: a description of the transgression (such as brazen, reckless), followed by a coding of past incidents fitting the description to create a baseline, followed by the addition of new incidents. Deep dives to analyze data for specific transgressions by specific adversaries helps to provide critical context and to differentiate between competing hypotheses. The overall results can be compared to deterrent policies and actions (as well as other, non-cyber developments between the nations) to see any suggestions of correlation.

Still, the interaction of international affairs and cyberspace is hidden, complex and

37 Perlroth, Nicole. 2010. "In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back." The New York Times, October 24. https://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html.

38 Harknett, Richard J. 2018. "United States Cyber Command's New Vision: What It Entails and Why It Matters." Lawfare, March 23. https://www.lawfareblog.com/united-states-cyber-commands-new-vision-what-it-entails-and-why-it-matters.

39 Mandiant. 2013. "APT1: Exposing One of China's Cyber Espionage Units." p4. https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf.

ever-changing, challenging the disentanglement of multiple causes and effects. Any methodology to measure policy impacts – not just the one presented here – will share the following shortcomings, each of which can be effectively minimized.

The most obvious shortcoming is that the effect of U.S. actions may be swamped by technical developments. An increase in the number of reported incidents could be due to new classes of vulnerabilities, a flood of new and insecure Internet-of-things devices, or improvements in detection and defense. The deployment of more secure infrastructure would lead to fewer attacks, as would an increase in adversary use of "living off the land" and obfuscation techniques. This class of shortcoming can be controlled for by assessments and metrics from cyber threat analysts directly tracking adversary operations (such as those following Chinese espionage before and after the Obama-Xi agreement). Any variances can be investigated by comparing against the trend lines of different adversaries (if, say, the Iran trend declines but the trend for China increases). Competing hypotheses ("this increase means little because of more deployment of insecure IoT devices") can be compared against actual observed adversary behavior.

A second set of shortcomings include that many attacks (and adversary motivations) are hidden and data can be hard to come by and analyze. Geopolitical events could cause adversaries to decrease or increase their use of cyber capabilities for strategic ends, regardless of U.S. counter-offensive operations. Fortunately, having an exact enumeration of the events in each category matters less than the direction and magnitude of the trends.

The advocates of persistent engagement and deterrence suggest it should have a substantial, perhaps unprecedented impact on adversary behavior. Anything other than a correspondingly strong reduction, such as that seen after the Obama-Xi agreement, suggests that the policy may not be working as intended. If the trend significantly worsens, it may be that a hypothesis that the new policy is inciting adversaries is a better fit to the curve. But it could also mean that any deterrent effect is being swamped by other signals, perhaps an overall rise in global incidents or a significant worsening of tensions with the adversary nation. Either of these can be checked against a control, such as the overall trend of global incidents and bilateral relations (such as US-China). Other controls can include target states (if the United States sees a decrease of brazen attacks from China while the United Kingdom and France see increases).

These shortcomings can also be addressed by more sharing of intelligence, assessments, and data sets. Different communities have different strengths. Academic researchers generally can only rely on open-source material, especially media reporting, but bring rigor and strict methodologies; while commercial cyber threat analysts have long

continuity following targets and have deep access to proprietary data (as the FireEye team did for the report on Chinese commercial espionage). U.S. government analysts, especially those in the Intelligence Community, can rely on classified sources but will miss much of the data held by commercial threat analysts (or by states, cities, and counties) and can overlook information not coming from classified sources.

A third set of shortcomings deal with methodological factors. The timescale to discover cyber incidents hampers assessment, as incidents are often not publicized until well after they are conducted, complicating efforts to ascribe cause and effect.[40] There may be so few truly dangerous attacks on a regular basis that an increase or decrease of a small number of incidents leads to an enormous percentage increase or decrease. These can be dealt with through appropriate structuring of the framework and coding of the data.

4. CONCLUSION AND FUTURE WORK

According to Michael Daniel, former White House cyber coordinator, the Trump administration "is willing to take more risks than previous administrations, but the proof will be in the results".[41] We can't assess what we don't try to measure. Together, the frameworks in this paper can act as a check on whether these new, riskier U.S. cyber policies and operations are succeeding in suppressing incoming attacks, or inciting them.

The shortcomings in the previous section are generally not specific to this paper and would pertain to *any* attempt at measuring the new U.S. policies. Some of the people reviewing this paper suggested that the U.S. Government – especially the intelligence community – would be uniquely placed to conduct these assessments. But the Federal Government cannot easily measure attacks from adversaries, as it lacks access to most victim data, which can be held by cybersecurity companies and organizations like the Cyber Threat Alliance. Moreover, the U.S. Government cannot easily even know all its own operations against adversaries: some will be covert actions, others espionage, while others are "traditional military operations." Each is held in a separate compartment and few individuals have the full picture.

Of course, we still encourage all parties to attempt to measure. The U.S. Government should conduct its assessment, with different agencies using their own processes,

40 This delay can be seen in large data breaches, such as the Marriott incident that occurred around the same time as the OPM incident in 2014 but was only publicized in 2018; as well as offensive cyber effects operations, such as U.S. Cyber Command reportedly blocking internet access to the Internet Research Agency on the day of the 2018 elections, which went unreported until February 2019.

41 Nakashima, Ellen. 2018. "White House authorizes 'offensive cyber operations' to deter foreign adversaries." The Washington Post, September 20. https://www.washingtonpost.com/world/national-security/trump-authorizes-offensive-cyber-operations-to-deter-foreign-adversaries-bolton-says/2018/09/20/b5880578-bd0b-11e8-b7d2-0773aa1e33da_story.html?utm_term=.f0d9d4720f36.

sources, and methods. The National Intelligence Council or Cyber Threat Intelligence Integration Center may be natural homes for much of this activity. As no one should be allowed to grade their own homework, this process should not be owned by either the National Security Agency or U.S. Cyber Command. The National Security Council must be the ultimate arbiter, deciding if the new operations are meeting the goals set by policymakers.

Rough-and-ready metrics such as those presented here can at least begin to indicate the direction and magnitude of changes over time, allowing indirect measurement to determine whether the policies are suppressing adversary attacks or inciting them. As this project moves forward, we will seek to improve and further refine the framework presented here for a usable pilot project for the commercial cyber threat intelligence community. These companies regularly assess the impact and quantity of foreign cyber operations; with a more standard and transparent methodology, they can help create a public understanding of the impact of U.S. actions on cyberspace, which has taken a central position in supporting our economy and society.

Additional research should also be done on historical antecedents of persistent engagement. Though the comparisons are inexact, persistent engagement has similarities to other examples where the military and intelligence forces of the two blocs during the Cold War were in routine belligerent contact: anti-submarine warfare; espionage-counterespionage; freedom of navigation operations; and intelligence, surveillance, and "exciter" flights against each other's homelands.

Acknowledgments

The authors wish to acknowledge Jennifer Gennaro of Columbia University's School of International and Public Affairs for her assistance. We would also like to thank the staff and attendees of the inaugural CyberWarCon conference (held on 28 November 2016 in Arlington, Virginia), where we presented an early version of this work; and the participants of a workshop at SIPA on 25 February 2019. Michael Daniel, Gregory Rattray, Chris Painter, and Alexandra Friedman also provided valuable support and input. This work was funded in part by the Office of Naval Research under the OSD Minerva program: Grant number N00014-17-1-2423

REFERENCES

Barnes, Julian E. 2018. "U.S. Begins First Cyberoperation Against Russia Aimed at Protecting Elections." *The New York Times*, October 23. https://www.nytimes.com/2018/10/23/us/politics/russian-hacking-usa-cyber-command.html.

Bolton, John. 2018. “Transcript: White House Press Briefing on National Cyber Strategy - Sept. 20, 2018”. Washington DC (September 8). Available at https://news.grabien.com/making-transcript-white-house-press-briefing-national-cyber-strateg.

Clarke, Richard A, and Robert K Knake. 2011. *Cyber War: The Next Threat to National Security and What to Do About It*. Ecco.

Constantine, Lucian. 2014. “New Havex Malware Variants Target Industrial Control System and SCADA Users.” *PC World*, June 24. www.pcworld.com/article/2367240/new-havex-malware-variants-target-industrial-control-system-and-scada-users.html.

Daniel, Michael. 2019. Interview with Michael Daniel (January 8).

Defense Science Board, Department of Defense. 2017. “Task Force on Cyber Deterrence.” Defense Science Board, 3, 4. https://www.acq.osd.mil/dsb/reports/2010s/DSB-cyberDeterrenceReport_02-28-17_Final.pdf.

Department of Defense. 2018. “Cyber Strategy 2018.” https://media.defense.gov/2018/Sep/18/2002041658/-1/-1/1/CYBER_STRATEGY_SUMMARY_FINAL.PDF.

Department of State. 2018. “Recommendations to the President on Deterring Adversaries and Better Protecting the American People from Cyber Threats.” May 31. https://www.state.gov/documents/organization/282253.pdf.

FireEye iSight Intelligence. 2016. “Redline Drawn: China Recalculated Its Use of Cyber Espionage.” FireEye. https://www.fireeye.com/content/dam/fireeye-www/current-threats/pdfs/rpt-china-espionage.pdf.

F-Secure Labs. 2014. “BlackEnergy & Quedagh: The Convergence of Crimeware and APT Attacks.” https://www.f-secure.com/documents/996508/1030745/blackenergy_whitepaper.pdf.

Geller, Eric. 2018. “Trump Scraps Obama Rules on Cyber Attacks, Giving Military Freer Hand.” *Politico*, August 18. https://www.politico.com/story/2018/08/16/trump-cybersecurity-cyberattack-hacking-military-742095.

Graff, Garrett M. 2018. “How the US Forced China to Quit Stealing—Using a Chinese Spy.” *Wired*, October 11. https://www.wired.com/story/us-china-cybertheft-su-bin/.

Harknett, Richard J. 2018. “United States Cyber Command’s New Vision: What It Entails and Why It Matters.” *Lawfare*, March 23. https://www.lawfareblog.com/united-states-cyber-commands-new-vision-what-it-entails-and-why-it-matters.

Healey, Jason. 2018. “Not The Cyber Deterrence the United States Wants.” *Council on Foreign Relations*, June 11. https://www.cfr.org/blog/not-cyber-deterrence-united-states-wants.

Isikoff, Michael, and David Corn. 2018. “‘Stand Down’: How The Obama Team Blew The Response To Russian Meddling.” *Huffington Post*, March 9. https://www.huffingtonpost.com/entry/stand-down-how-the-obama-team-blew-the-response-to-russian-meddling_us_5aa29a97e4b086698a9d1112.

Jackson Higgins, Kelly. 2016. “Lessons from the Ukraine Electric Grid Hack.” *Dark Reading*, March 18. https://www.darkreading.com/vulnerabilities---threats/lessons-from-the-ukraine-electric-grid-hack/d/d-id/1324743.

Kissinger, Henry. 1994. *Diplomacy*. Simon and Schuster.

Langevin, James R. 2019. “Opening Statement: FY 2020 Budget Request for Military Operations in Cyberspace.” March 13. https://armedservices.house.gov/_cache/files/d/5/d5f94725-3373-40ef-803c-1f0ff8f106a8/577D710BF48F37825B2656EE1AF6891A.opening-statement---ietc-chairman-langevin-3-13-2019.pdf.

Mandiant. 2014. “APT1: Exposing One of China’s Cyber Espionage Units.”

Marczak, Bill, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. 2015. "China's Great Cannon." *The Citizen Lab*. April 10. https://citizenlab.ca/2015/04/chinas-great-cannon/.

Microsoft. 2018. Digital Peace. https://digitalpeace.microsoft.com/.

Nakashima, Ellen. 2014. "U.S. Rallied Multinational Response to 2012 Cyberattack on American Banks." *The Washington Post*, April 11. https://www.washingtonpost.com/world/national-security/us-rallied-multi-nation-response-to-2012-cyberattack-on-american-banks/2014/04/11/7c1fbb12-b45c-11e3-8cb6-284052554d74_story.html?utm_term=.be386c400c97.

—. 2018. "White House authorizes 'offensive cyber operations' to deter foreign adversaries." *The Washington Post*, September 20. https://www.washingtonpost.com/world/national-security/trump-authorizes-offensive-cyber-operations-to-deter-foreign-adversaries-bolton-says/2018/09/20/b5880578-bd0b-11e8-b7d2-0773aa1e33da_story.html?utm_term=.f0d9d4720f36.

Nakasone, Paul M. 2019. "An Interview with Paul M. Nakasone." *Joint Forces Quarterly*. https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-92/jfq-92.pdf.

Office of the Director of National Intelligence. 2017. "Background to 'Assessing Russian Activities and Intentions in Recent US Elections': The Analytic Process and Cyber Incident Attribution." https://www.dni.gov/files/documents/ICA_2017_01.pdf.

Oxford. 2009. "Recklessness." In *Oxford English Dictionary*. Oxford University Press.

Paris Peace Forum. n.d. Paris Peace Forum. https://parispeaceforum.org/.

Perlroth, Nicole. 2010. "In Cyberattack on Saudi Firm, U.S. Sees Iran Firing Back." *The New York Times*, October 24. https://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html.

Perlroth, Nicole, and Quentin Hardy. 2013. "Bank Hacking Was the Work of Iranians, Officials Say." *The New York Times*, January 8. https://www.nytimes.com/2013/01/09/technology/online-banking-attacks-were-work-of-iran-us-officials-say.html.

Reuters. 2018. "US Accuses China of Violating Bilateral Anti-Hacking Agreement." CNBC, November 8. https://www.cnbc.com/2018/11/09/us-accuses-china-of-violating-bilateral-anti-hacking-agreement.html.

Sanger, David E. 2018. *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age*. Crown.

Segal, Adam. 2016. "The U.S.-China Cyber Espionage Deal One Year Later." *Council on Foreign Relations*, September 28. https://www.cfr.org/blog/us-china-cyber-espionage-deal-one-year-later.

Sulmeyer, Michael. 2018. "How the U.S. can Play Cyber-Offense." *Foreign Affairs*, March 22. https://www.foreignaffairs.com/articles/world/2018-03-22/how-us-can-play-cyber-offense.

The White House. 2018. "National Cyber Strategy." https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf.

U.S. Cyber Command. 2018. "Achieve and Maintain Cyberspace Superiority: Command VIsion for U.S. Cyber Command." April. https://www.cybercom.mil/Portals/56/Documents/USCYBERCOM%20Vision%20April%202018.pdf?ver=2018-06-14-152556-010.

U.S.-CERT, Department of Homeland Security. n.d. "NCCIC Cyber Incident Scoring System." https://www.us-cert.gov/NCCIC-Cyber-Incident-Scoring-System.

—. n.d. "US-CERT Federal Incident Notification Guidelines." https://www.us-cert.gov/incident-notification-guidelines.

Valeriano, Brandon, and Ryan C Maness. 2015. *Cyber War versus Cyber Realities. Cyber Conflict in the International System*. Oxford University Press.

Zetter, Kim. 2014. "An Unprecedented Look at Stuxnet, the World's First Digital Weapon." *Wired*, November 3. https://www.wired.com/2014/11/countdown-to-zero-day-stuxnet/.