



## A tale of two cybers - how threat reporting by cybersecurity firms systematically underrepresents threats to civil society

Lennart Maschmeyer , Ronald J. Deibert & Jon R. Lindsay

To cite this article: Lennart Maschmeyer , Ronald J. Deibert & Jon R. Lindsay (2020): A tale of two cybers - how threat reporting by cybersecurity firms systematically underrepresents threats to civil society, Journal of Information Technology & Politics, DOI: [10.1080/19331681.2020.1776658](https://doi.org/10.1080/19331681.2020.1776658)

To link to this article: <https://doi.org/10.1080/19331681.2020.1776658>



© 2020 The Author(s). Published with license by Taylor & Francis Group, LLC



Published online: 11 Jun 2020.



[Submit your article to this journal](#)



Article views: 9921



[View related articles](#)




[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

## A tale of two cybers - how threat reporting by cybersecurity firms systematically underrepresents threats to civil society

Lennart Maschmeyer , Ronald J. Deibert, and Jon R. Lindsay

### ABSTRACT

Public and academic knowledge of cyber conflict relies heavily on data from commercial threat reporting. There are reasons to be concerned that these data provide a distorted view of cyber threat activity. Commercial cybersecurity firms only focus on a subset of the universe of threats, and they only report publicly on a subset of the subset. High end threats to high-profile victims are prioritized in commercial reporting while threats to civil society organizations, which lack the resources to pay for high-end cyber defense, tend to be neglected or entirely bracketed. This selection bias not only hampers scholarship on cybersecurity but also has concerning consequences for democracy. We present and analyze an original dataset of available public reporting by the private sector together with independent research centers. We also present three case studies tracing reporting patterns on a cyber operation targeting civil society. Our findings confirm the neglect of civil society threats, supporting the hypothesis that commercial interests of firms will produce a systematic bias in reporting, which functions as much as advertising as intelligence. The result is a truncated sample of cyber conflict that underrepresents civil society targeting and distorts academic debate as well as public policy.

### KEYWORDS

Cybersecurity; public goods; civil society; threat intelligence; cyber conflict; market failure

On October 1, 2018, a Citizen Lab report revealed that the phone of Omar Abdulaziz, a prominent dissident of the Kingdom of Saudi Arabia, had been infected with sophisticated spyware (Marczak et al. 2015). The researchers established with a high degree of confidence that his phone was compromised by an operator associated with the Saudi Arabian government; they also identified the spyware as the ‘Pegasus’ suite manufactured by the Israel-based vendor NSO Group. Abdulaziz, a university student and Canadian resident, runs a popular YouTube channel posting satirical videos critical of the Saudi regime. One day later, another high-profile dissident, Washington Post journalist Jamal Khashoggi, was lured into the Saudi consulate in Istanbul, Turkey, where he was murdered and dismembered. Soon thereafter it was revealed that Abdulaziz and Khashoggi were working together on a social media opposition campaign against the Crown Prince of Saudi Arabia, principally communicating over an encrypted, supposedly private WhatsApp conversation but which Citizen Lab discovered was being remotely monitored by Saudi intelligence. Although the reasons underlying the specific decision to murder

Khashoggi are unknown, many have drawn connections to the surveillance uncovered in this operation (Rogin 2018; Shezaf 2018).

This case calls into question several popular assumptions about cybersecurity. First, prevalent narratives emphasize threats to critical infrastructure, intellectual property, and state secrets. In this case, however, Saudi Arabia used a sophisticated exploitation platform to target a lone critic running a comedy channel. Second, cyberspace is widely thought to advantage weaker actors against the strong, but the asymmetry here runs in the reverse direction.<sup>1</sup> NSO Group, a self-described “cyber warfare” firm valued at 1 billion USD, sells its technology exclusively to government, law enforcement, military, and intelligence agencies, yet the Saudi regime used its military-grade capabilities to hack a dissident’s iPhone. Third, security firms and government agencies are usually considered experts in cybersecurity, yet this threat was identified and disclosed by civil society itself. Citizen Lab is a small unit of a major research university that conducts interdisciplinary research into targeted digital threats,<sup>2</sup> and it detected the threat to Abdulaziz only because it was studying a broader pattern of

**Table 1.** Summary of case studies.

Cases	TTP	Threat actor	Victim profile	Case type
Tainted Leaks	Unique, high sophistication	High-profile (APT28, Russia)	High-profile (journalist + DNC, Hillary Clinton)	Most-likely
Spying on a Budget	Low sophistication	Low-profile(unknown actor)	Low- profile (Tibetan activists)	Most-likely
Familiar Feeling	Medium sophistication	Medium-profile (TropicTrooper, speculated link to China)	Medium-profile (Tibetan activists + governments)	Least-likely

human rights violations abetted by NSO Group. While scholarly and policy discourse about cybersecurity focuses on high-end threats to high-profile actors, there are reasons to believe that the targeted exploitation of civil society is a fundamental feature of the cyber revolution. Russian interference in the 2016 U.S. presidential election dramatically highlighted the vulnerability of civil society and the shortcomings of the cybersecurity debate, precisely because the victim was a democratic superpower. More typically, civil society has suffered in silence.

What explains the gap between perception and practice? Most of what we know about cyber conflict stems from data provided by public reporting from cybersecurity firms, yet in these reports civil society targets like Omar Abdulaziz tend to receive only passing attention. We consider commercial threat reporting in terms of the publicly available reports on targeted digital threats by private vendors of threat intelligence services. These reports are primarily marketing instruments aiming to increase revenue from the paid products offered by these vendors, namely private intelligence reporting and network defense services. Importantly, the incentives driving public reporting tend to create a biased sample of incidents at the high end of conflict spectrum and/or targeting rich actors who can afford to pay for commercial cyber defense. Threats against civil society organizations (CSOs)<sup>3</sup> who cannot afford to pay, however, tend to go unreported while their networks go undefended. This is bad for both the health of democracy and the study of cybersecurity.

Hence, we argue, commercial threat reporting presents a truncated sample of cyber conflict that distorts threat perceptions. This reporting is subject to systematic bias, yet this bias has not been systematically examined. Threat inflation within commercial reports is a well-established problem, but selection bias *across* reports has not been sufficiently addressed in the literature on cyber conflict.

Importantly, because commercial threat reporting offers by far the largest, and often the only, source of data on cyber conflict, this bias is likely to impact perception among both policy-makers and researchers.

We test this theory with an original dataset of all available public reporting on targeted exploitation, comprised of 700 reports in total, from 2009–2018. The reports we collected were derived from two types of sources: first, commercial threat intelligence vendors (629 reports), and second, independent research centers (71 reports). We also examine helpline data from AccessNow, a digital rights advocacy group, reflecting digital threats as reported by civil society itself. We find that a low proportion of commercial threat reports discuss civil society, and those that do focus on high-profile victims and threat actors. The geographical distribution of reporting and attribution patterns are congruent with the hypothesized selection bias. As a further plausibility probe, we select three cases of civil society exploitation, one attributed to Russia and the other two to China, for structured focused comparison. The forensic data strongly confirm our theory as even the least-likely case exhibits clearly selective reporting. We conclude with a discussion of implications for scholarship and democracy in the digital age.

### The problem: what do we know?

Early cybersecurity scholarship imagined a future of cyberwar that relied mostly on speculation since data was scarce. John Arquila and David Ronfeldt hypothesized in 1993 that “the information revolution implies the rise of cyberwar, in which neither mass nor mobility will decide outcomes” (1993, p. 141). In this scenario, information trumps mass, geography is secondary as conflict occurs ‘in cyberspace’ at unprecedented speeds, and power is diffused toward smaller actors, leading to a rise of

asymmetric threats (Arquilla & Ronfeldt, 1993, p. 143–55). This cyberwar scenario became conventional wisdom in cybersecurity discourse and persisted for almost two decades.

Consequently, in 2010 Lynn III argued that “cyberwarfare is asymmetric . . . . A dozen determined computer programmers can, if they find a vulnerability to exploit, threaten the United States” (Lynn, 2010). Similarly, Joseph Nye asserts that “low barriers to entry contribute to the diffusion of power in the cyberdomain” (Nye, 2011, p. 124). This perception led to a focus on threats to critical infrastructure, and the assumption that vulnerabilities in the latter “provide asymmetrical advantages to nonstate actors” (Rattray, 2009, p. 265). Yet the threat of destructive cyberwar has remained hypothetical, and even the best candidate for an exception (Stuxnet) violates conventional assumptions about cyberwar (Lindsay, 2013), thus proving the rule.

Around the same time, however, Citizen Lab demonstrated that empirical data collection on cyber conflict was possible as its researchers discovered that a hacking operation against the Tibetan exile government was part of a global Chinese espionage campaign targeting government agencies and civil society across 103 countries (R. Deibert & Rohozinski, 2009). This report, titled “Ghost Net”, changed the cybersecurity landscape, as private vendors started publishing reports under names like “Shady RAT” and “Aurora” detailing cyber operations discovered in the wild (HBGary, 2010; McAfee, 2010). The same year the Stuxnet malware offered the first evidence of cyber operations causing physical damage. Importantly, most empirical evidence on this operation came from commercial threat reports (ESET, 2010; Langner, 2011; Symantec, 2011). Such public reporting, freely available on vendors’ websites, quickly increased in volume, providing both scholars and policy-makers with a rich new source of data.

Building on these data, several scholars challenged the established wisdom on cyberwar. In 2012, Thomas Rid argued that the “the world never experienced an act of cyber war . . . instead, the last decade saw increasingly sophisticated acts of network-enabled sabotage, espionage, and subversion.” (Rid, 2012, p. 29). Jon Lindsay’s study of Stuxnet (2013), primarily relying on commercial

threat reporting data, underlined that “the technically and organizationally sophisticated level of play required for cyber warfare is generally beyond the capacity of a lone hacker, a small group of amateurs, or even organized criminals” (Lindsay, 2013, p. 389). Similarly, Erik Gartzke concludes that “by far the most compelling scenario for the transformation of political conflict through the internet . . . involves the use of the internet for espionage” (Gartzke, 2013, p. 70). The limited effectiveness of cyber operations as means of warfare, coercion and destruction has since been emphasized by an increasing number of scholars (Borghard & Lonergan, 2017; Slayton, 2017; Smeets, 2018). In short, acts of violence or physical destruction involve non-trivial operational challenges that only powerful actors are both likely to overcome and unlikely to have interests in, given other more reliable ways of generating coercive harm. By contrast, digital espionage offers significant gains at relatively low risk.

Meanwhile, the volume of commercial threat reporting has increased exponentially and academics increasingly rely on the rich data these reports provide. Examples include Lindsay’s survey of Chinese cyber espionage (2014), Thomas Rid and Ben Buchanan’s analysis of attribution (2015), Buchanan’s 2017 book *The Cybersecurity Dilemma*, and Ben Jensen, Ryan Maness and Brandon Valeriano’s case study of Russia (2019). This growing availability of publicly accessible data is clearly a positive development.

Yet commercial actors adopted the method pioneered by Citizen Lab, but not the substance of reporting. Commercial threat reporting primarily focuses on cybercrime, economic espionage and sabotage of critical infrastructure (CrowdStrike, 2019, 2; FireEye, 2019, 4; Symantec, 2019, p. 2–3). Since these reports constitute the largest, and often the only, source of data on cyber operations, perceptions of both policy-makers and academics can be expected to reflect patterns evident in threat reporting. Accordingly, JD Work underlines that due to the lack of alternative sources of data, “policymakers, military professionals, and scholars must rely heavily on this new range of sources to understand developments in the cyber domain” (Work, 2020, p. 2). Consequently, “a growing number of major policy issues are . . . profoundly shaped

by underlying commercial intelligence reporting” (Work, 2020, p. 2). What is reported and what is not thus has a significant influence on academia and policy.<sup>4</sup>

The interference in the 2016 US Presidential elections through leaked information and social media influence campaigns upset prevailing threat models. As Jayamaha and Matissek put it, “no one expected that ‘subversive instruments’ would be used in such a way as to create intra-societal tensions through exploitation of civil society organizations” (2018). Subsequent investigations – including by special counsel Robert Mueller III (2019) – revealed a large-scale influence campaign using disinformation to sway voter opinions and foster divisions (Isaac & Wakabayashi, 2017; ODNI, 2017). Its actual effects on election outcomes continue to be hotly debated, but its significance is clear and reflected in threat perceptions of this operation as an ‘act of war’ (Schleifer & Walsh, 2017). This Russian influence campaign focusing on individuals and civil society caught most scholars and policy-makers off guard; it did not correspond to prevailing threat models focusing on critical infrastructure disruption and large-scale digital espionage.

The collective surprise among scholars and policy-makers alike suggests commercial threat reporting, a key data source informing prevailing threat models, provides an incomplete picture of cyber conflict. Significantly, in stark contrast to Jayamaha and Matissek’s claimed surprise, one scholar warned as early as 2003 that “pressures from the security and commercial sectors to regulate and control the Internet are beginning to alter its basic material architecture in ways that may undermine not only the activities of global civic networks, but also the long-term prospects for an open global communications environment” (R. J. Deibert, 2003). Others have noted the threat cyber conflict poses to civil society (Brantly, 2014), yet in academia it has remained mostly a fringe topic. Meanwhile, independent research centers and nonprofit organizations have documented the proliferation of targeted digital threats to civil society for over a decade.<sup>5</sup> Independent research centers have only a fraction of the resources of commercial vendors, however, limiting their

capacity to investigate and report on such threats at scale.

Threat intelligence has become a multibillion-dollar industry, hence vendors have the resources to report on many different threat types. However, commercial threats reports are unlikely to provide a representative sample of cyber conflict due to underlying business incentives. Existing research has addressed shortcomings of current attribution processes (Egloff, 2020; M. Mueller et al., 2019) and a trend of threat inflation in commercial reports to increase sales of security products (Dunn Cavely, 2013). Since commercial reporting is usually the main source of data on cyber conflict, this threat inflation likely shaped exaggerated fears of ‘cyber doom’ (Lawson, 2013) and cyber terrorism that distort current debates and policy-making (Myriam Dunn-Cavelty, 2008). Yet such problems within reporting is well-established, systematic bias *across* reporting by different firms, affecting what types of threats are reported, and which are not, has not been sufficiently explored. A closer examination of the incentives behind threat reporting lead us to expect not only threat inflation, but systematic underreporting of specific threat types – and particularly threats to civil society. This problem is especially acute because of the scarcity of alternative data sources, and the dire consequences for civil society itself.

### **Commercial threat reporting presents a truncated sample of cyber conflict**

We argue that commercial threat reporting presents a truncated sample of cyber conflict due to the private interests that shape reporting. Profit incentives lead firms to prioritize high-end threats to powerful actors using unique methods in their reporting, while neglecting threats to weaker actors – in particular, civil society. Consequently, threat reporting provides a distorted understanding of targeted cyber threats that focuses on activity at the high end of the conflict spectrum, and neglecting or bracketing activity at the lower end. This situation constitutes a market failure that leaves those most in need of accurate information about threats – vulnerable civil society actors – least well-informed. Moreover, because commercial reporting is often the only source of data, the distorted

perception it provides to policy makers and academics results in an under prioritization of the problem.

Commercial threat reporting is part of the threat intelligence and network defense sector in information security. Threat intelligence firms are profit-driven enterprises that generate three main products: freely available public reporting, more comprehensive private reporting available to paying subscribers, and custom protection services that come at a substantial premium. We focus on public reports, where private vendors publish findings of their investigations into cyber operations detailing the tactics, techniques and procedures (TTP) used by so-called ‘threat actors’ to breach systems for data theft, surveillance and/or disruption.<sup>6</sup>

Importantly, public reporting is foremostly a marketing instrument to increase revenue from the two premium services mentioned above. Cybersecurity, in terms of both threat magnitude and defensive effectiveness, is notoriously hard to measure (Anderson et al., 2013). When a firm cannot directly advertise its comparative advantage, it will resort to indirect measures. A technically detailed report on a dramatic intrusion into a high-value target says, in effect, because our employees are smart enough to reverse engineer cyberwarfare, they are also smart enough to protect your business from it. According to Juan Andrés Guerrero-Saade, then a researcher at prominent vendor Kaspersky, “the intended purpose is a PR-coup to both attract new customers for closed-release intelligence reports as well as garner brand recognition and industry respect for formidable findings” (Guerrero-Saade, 2015, p. 4). Similarly, JD Work underlines that a ‘majority’ of public reporting constitutes ‘marketing collateral’ designed to “attract new customers, position themselves for evaluation by industry market research analysts” and “engage with prospective investors,” (Work, 2020, p. 16).

Consequently, commercial reports typically have two parts: first, they inform the audience about threats, and second, they highlight products to alleviate these threats. Vendors can be expected to target those sectors most likely to buy their products. CSOs, however, are notoriously cash-strapped (Crete-Nishihata et al., 2014, 2; CLTC, 2018), and thus least likely to invest in premium security products. In short, sectors of the greatest

interests to threat intelligence vendors – government, military, Fortune-500 firms, etc – are likely to be prioritized in reporting, while low-revenue sectors are likely to be neglected or entirely ignored. If these assumptions are right, threat reporting presents a distorted picture of cyber conflict where threats aligned with the profit incentives of cybersecurity vendors are overrepresented, while civil society threats are underreported or entirely missing. This expectation aligns with Egloff’s contestation that only financially potent or politically relevant targets “have the public visibility for security companies to show off their skills” (Egloff, 2020, p. 7). The result is a classic market failure, i.e., the “failure of a more or less idealized system of price-market institutions to sustain ‘desirable’ activities or to stop ‘undesirable’ activities” (Bator, 1958, p. 351).

Although concentrated business interests shape what goes into public reporting, the product nonetheless provides diffuse benefits to the wider cybersecurity community about targeted threats.<sup>7</sup> Accordingly, Rosenzweig has proposed characterizing commercial threat reporting as a public good (2011). Threat reports fulfil an important role because the knowledge they provide not only helps scholars better understand cyber conflict, but it is essential for practitioners and potential victims to increase resilience and mitigate intrusions. Nascent community-driven initiatives to consolidate knowledge from threat reporting in shared resources attests to this importance.<sup>8</sup>

The underprovision of threat reporting to CSOs that results from the profit-incentives driving it has two key consequences. First, these organizations lack information on the threats they face. CSOs generally lack technical expertise and resources, making them easy targets. Moreover, they are also attractive targets for security services interested in surveilling, exploiting, or repressing them. Potential consequences are more severe compared to commercial actors because they involve personal harm, detention, or even death (Crete-Nishihata et al., 2014, p. 117). Therefore, CSOs urgently need accurate threat information. Second, underreporting of threats to civil society exacerbates their lack of defenses because it leads to insufficient prioritization of the issue by both policy-makers and funders.

### Sources of bias in reporting

Commercial reporting is driven by specific business interests that determine what gets reported, and what does not. The resulting selection criteria can be expected to produce a truncated sample of cyber conflict. As marketing instruments, threat reports need to maximize attention. Based on public statements, existing research and one formal interview with a threat intelligence researcher at a prominent firm, we identify three key selection criteria that shape reporting. Threats to civil society tend to score low across all three and can thus be expected to be neglected or entirely bracketed in commercial threat reporting.

First, a cyber operation exhibits some *unique characteristics*, typically in its TTP. According to a threat intelligence researcher, to make it into a public report “it needs to be something unique, something that hasn’t been reported before, for example a zero-day, or some kind of unique tactic used” (Threat Intelligence Researcher 2018).

Second, it has a *high-profile victim*. Since threat reports are intended to sell protection products, the more significant the threat, and the more high-profile<sup>9</sup> the targeted actor, the better. From the perspective of threat intelligence firms, the highest profile actors are those with the greatest revenue potential. If threat reporting is intended to sell private reports and protection services, a rational profit-seeking actor can be expected to prioritize reporting on threats to the most lucrative targets. Since the Global North is more affluent, and since most firms are headquartered in the Global North, facilitating sales, reporting can be expected to prioritize threats targeting this region

Third, a *high-profile threat actor* is behind the campaign. We identify three key measures of high-profile actors: (1) attribution to strategic competitors of nation-state(s) in which the target audience resides; (2) previous coverage in general news outlets, and (3) attribution to previous campaigns perceived as a national or international threat. The majority of threat intelligence firms are based in North America (Kuerbis & Badiei, 2017, p. 471–72), hence threats by the main adversaries perceived by a North American audience – Russia, Iran, China and North Korea (YouGov, 2017) – can be expected to be prioritized.

We expect targeted threats to civil society score low on at least two of these three variables: first, due to porous defenses, attackers can often rely on generic and cheap methods; and, second, their lack of purchasing power renders CSOs unattractive clients – and thus low-profile actors from the perspective of threat intelligence vendors.<sup>10</sup> In conclusion, we expect commercial threat reporting to present a truncated sample of cyber conflict that distorts perceptions of the priorities and methods of capable threat actors. We expect it to prioritize the high-end of cyber conflict: high-profile actors going after high-profile targets with sophisticated and unique methods. If the selection criteria identified here are accurate, the low end, where most of the targeting of civil society occurs, will be either neglected or entirely bracketed.

### Hypotheses and research design

We hypothesize that business incentives result in systematic selection bias in threat reporting.<sup>11</sup> To test our assumptions, we employ a mixed method research design following Lieberman’s nested approach, which “combines the statistical analysis of a large sample of cases with the in-depth investigation of one or more of the cases contained within the large sample.” (Lieberman, 2005, p. 434–35). We proceed in three steps. First, we formulate a set of hypotheses. Second, we test them against summary statistics drawn from our dataset of all available threat reporting. Third, we proceed with a qualitative analysis of three case studies to verify whether our predictions are congruent with reporting patterns on these cases.

#### Part I: threat reporting data

We present a new dataset of all available public reporting on targeted threats and employ content analysis to identify overall reporting patterns. This approach leverages the strength of large-N analysis in identifying broad trends (Lieberman, 2005, p. 436). Content analysis is a useful tool because it allows quantitative analysis of unstructured data to identify trends and potential biases (Mukherjee, 2018, p. 29–30). Our main research question is straightforward: what threats are being reported by commercial vendors? Our auxiliary research

question is: how do reporting patterns differ among commercial and independent reporting?

Our dataset comprises 700 threat reports, 629 reports by threat intelligence firms and 71 reports by independent research centers. Reports were collected from vendor/organization websites and community-run resources. Selection criteria were straightforward: to be included, reports had to discuss (1) a targeted digital threat, and (2) be available publicly.<sup>12</sup> We then specified a set of categories and coded all reports across these categories, before using descriptive statistics to verify the following three hypotheses (capturing three types of selection bias)<sup>13</sup>:

*H1: Threats to civil society are underreported in commercial threat reports.*

To test H1, we rely on two indicators: first, the overall proportion of reporting on civil society, and second, reporting on commercial spyware. The hypothesis is confirmed if only a small proportion of reports discuss threats to civil society. Conversely, if a majority of reports discuss civil society threats, it is false. Concerning the second indicator, commercial reporting on the targeted use of commercial spyware against civil society is expected to be non-existent to minimal. Our hypothesis would be disconfirmed if the analysis instead shows a significant proportion of commercial reporting focusing on spyware.

*H2: Reporting is geographically skewed toward the Global North.*

We use geographical distribution as an indicator of high-profile targets because it is the most uniformly reported metric.<sup>14</sup> If commercial reporting exhibits the expected geographical skew, and independent reporting does not, it further corroborates our hypothesis. If both commercial and independent reports exhibit identical or similar geographical bias, the hypothesis is invalidated. The same applies to the comparison with self-reporting data provided by AccessNow.

*H3: Reporting is skewed toward operations attributed to the target audience's main adversaries.*

We chose attribution to measure the profile of a threat actor because threat reporting does not

provide data on the values for the other hypothesized determinants of a high-profile actor (public & media attention, previous campaigns).<sup>15</sup> Attributing cyber operations to governments is risky due to potential repercussions (Guerrero-Saade, 2015), hence firms often avoid conclusive statements on this matter. However, commercial reports provide a wealth of indicators and attribute operations to specific state actors with a reasonable degree of confidence in about half of the cases (49%) – allowing tracking of attribution patterns.<sup>16</sup>

If a majority of reporting discusses threats by the main strategic competitors of the United States (Russia, China, Iran, North Korea) the hypothesis is verified. If, however, reporting is evenly spread across different threat actors, or otherwise distributed, the hypothesis is invalidated. We compare attribution patterns across commercial and independent reporting to spot divergences.

To conclude, our aim is not to show that independent reporting is more representative than commercial reporting – it has its own selection bias prioritizing civil society. Instead, we simply aim to test our hypothesis that commercial reporting provides a truncated sample by showing that there is reason to believe that additional threat phenomena exist in the world that could be, but are not, reported by commercial firms. The quantitative analysis constitutes a hoop-test of our theory (Van Evera, 1997, p. 31), meaning negative findings eliminate the theory yet positive findings do not invalidate rival explanations.<sup>17</sup> Content analysis does not allow for confirming causal relationships among variables, however (Mukherjee, 2018, p. 36). Therefore, positive results of this analysis alone cannot confirm the presence of systematic bias in commercial threat reporting.

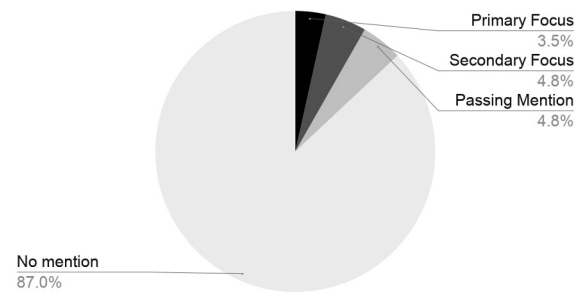
## **Part II: case studies**

Hence, we shift the level of analysis to qualitative analysis of three cases from within the same dataset (Lieberman, 2005, p. 440) that serve as a plausibility probe. We employ congruence testing and controlled comparison to “match the predictions and expectations of the theory with the outcomes of the cases to see if they are consistent” (George & Bennett, 2005, p. 227). The aim is to verify the predicted bias in commercial reporting toward



operations involving unique TTP, pursued by high-profile threat actors against high-profile victims. Each case involves a distinct cyber operation targeting civil society analyzed in a Citizen Lab report and involves three analytical steps corresponding to the three sources of selection bias identified. First, we track the overall volume of commercial reporting on the threat actor involved. Second, within this sample of reporting on the actor in question, we trace the prioritization of civil society targeting vis-à-vis other targets. Third, within the same sample, we trace prioritization of highly sophisticated TTP versus less sophisticated methods used by the same actor.

We select most- and least-likely cases covering the full range of values on each of the three selection criteria, constituting the independent variables. Our two most-likely cases are at the extreme ends of the spectrum, where our theory would predict a very high, and a very low, respectively, volume of reporting and prioritization within individual reports. The first involves a high-profile actor using unique TTP against a high-profile target, the other a low-profile actor using generic TTP against low-profile targets. The least-likely case involves a medium-profile actor targeting a medium-profile target with somewhat advanced TTP, where our theory does not make a strong prediction about reporting volume and prioritization of threats to civil society. If our predictions fail in the most-likely cases, strong doubt is

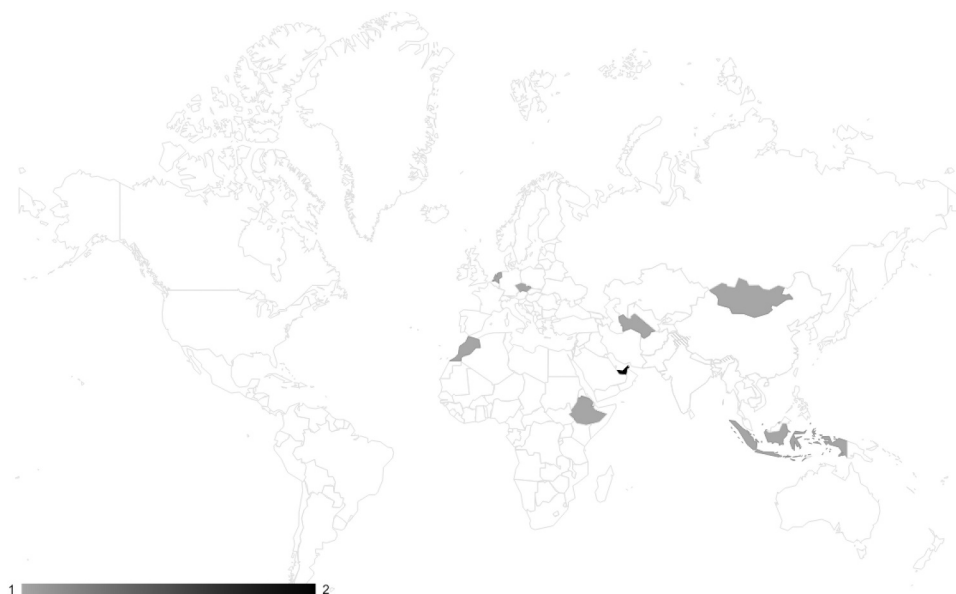


**Figure 1.** Commercial reporting – reporting volumes and prioritization of civil society targeting.

cast on our theory, while successful prediction in a least-likely case strongly supports the theory (George & Bennett, 2005, p. 147).

### Findings: content analysis

H1 predicts a low proportion of reporting prioritizing threats to civil society, and is verified by the findings. As shown in Figure 1, only a small minority, 82 out of the 629 commercial reports analyzed (13%), discuss a targeted threat to civil society. A deeper look at prioritization of the issue within this subset of commercial reporting revealed that only 22 out of these reports (4% of total reporting) place their primary focus on civil society. Meanwhile, 30 reports (5%) place a secondary focus on civil society targeting, with limited analysis, and 30 reports (5%) mention civil society in



**Figure 2.** Independent reporting – commercial spyware infrastructure detection 2012.

only passing. These results provide strong support for H1.

Findings on commercial spyware are also striking, confirming the expected neglect of the issue in commercial reporting.

The figure above shows spyware infrastructure as detected by Citizen Lab reports in 2012, comprising 12 countries. In 2018 this number had proliferated massively to 51 countries (Figure 3, below).

The growth of command and control infrastructure is an indication of the use of the associated spyware tool by the host government, but it does not reveal targets of such spyware (Marczak et al., 2015). However, our data shows that Citizen Lab reports have tracked the use of spyware against civil society in 22 of these countries – a third of the cases.

In comparison, only 8 out of 629 commercial threat reports (>1%) track the targeted use of commercial spyware, and two mention civil society targeting. This miniscule fraction of commercial reporting on commercial spyware, in contrast to its evident global proliferation, strongly confirms the hypothesized underreporting of threats to civil society.<sup>18</sup>

H2 predicts a geographical bias toward the Global North, which is supported by our findings on the absolute distribution of reporting, and further corroborated by the relative distribution as compared to independent reporting.

Figure 4 projects the geographical location of CSOs that have become victims of targeted digital threats in commercial reporting, counting the number of reports. The high concentration of reporting on targets in China (25) could be interpreted to challenge our hypothesis, yet this pattern corresponds to the hypothesized focus on strategic competitors (see below). Meanwhile, apart from three reported operations targeting civil society in Egypt, Africa and South America remain a blank spot.

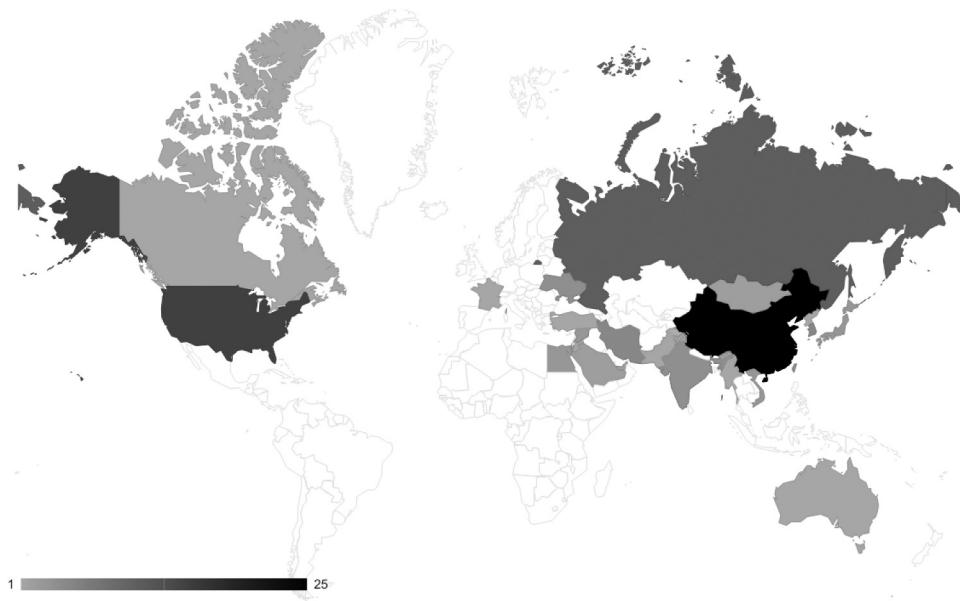
In contrast, independent reporting (Figure 5) reveals a host of targeted threats to civil society on these two continents. There are six instances in four different countries in South and Central America (Mexico, Ecuador, Brazil, Paraguay), and twelve reported instances in five different countries in Africa (Morocco, Egypt, Ethiopia, Kenya, South Africa). The absolute geographical distribution of commercial reporting on targeted threats to civil society thus supports H2, which is further supported by its relative distribution compared to independent reporting.

Comparison to AccessNow helpline data further underlines the underrepresentation of the Global South.

Although this sample only covers 17 months (January 2016 – May 2018), it shows a much wider and more even distribution than commercial



Figure 3. Independent reporting – commercial spyware infrastructure detection 2018.



**Figure 4.** Commercial reporting – civil society targeting by country.



**Figure 5.** Independent reporting – civil society targeting by country.

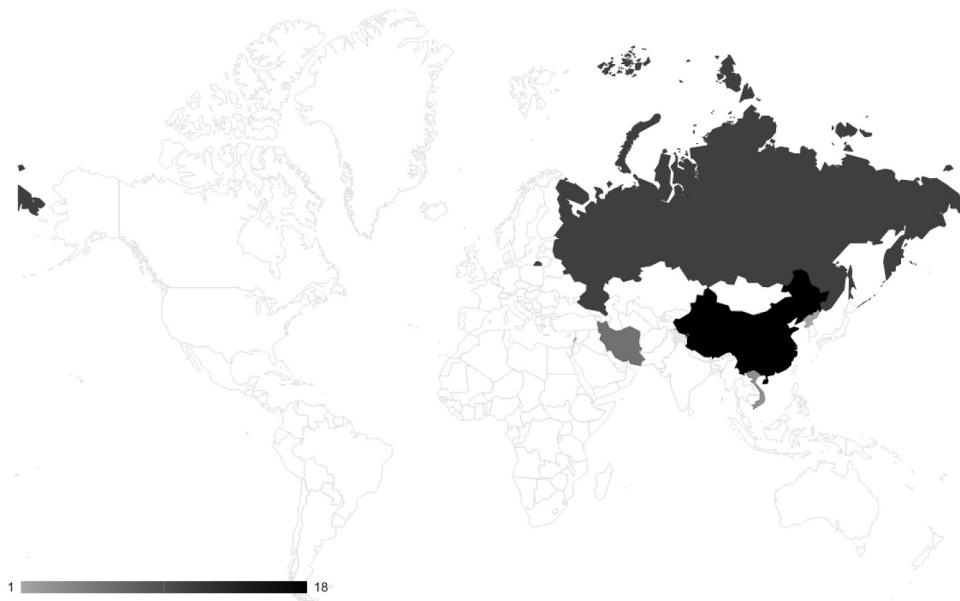
reporting (Figure 4). There is a plethora of activity in the Global South that is missing from commercial reporting. Granted, these are very different sources of data and thus not easily comparable – nonetheless, the divergence reinforces the trend identified in the comparison above, thus providing further support to our hypothesis.

H3 predicts reporting to be skewed toward perceived adversaries of the target audience. Our findings are strikingly clear.

As depicted in Figure 7, commercial reporting attributes the vast majority (88%) of targeted threats to civil society to the United States' key strategic competitors: China (18), Russia (11) and Iran (6). Only five campaigns are attributed to other states: Vietnam (3), North Korea (1) and Lebanon (1). This pattern is strikingly congruent with the hypothesized bias toward the perceived 'main adversaries' of a North American audience. Russia, China and Iran are often counted among the world's leading



**Figure 6.** AccessNow – helpline data.



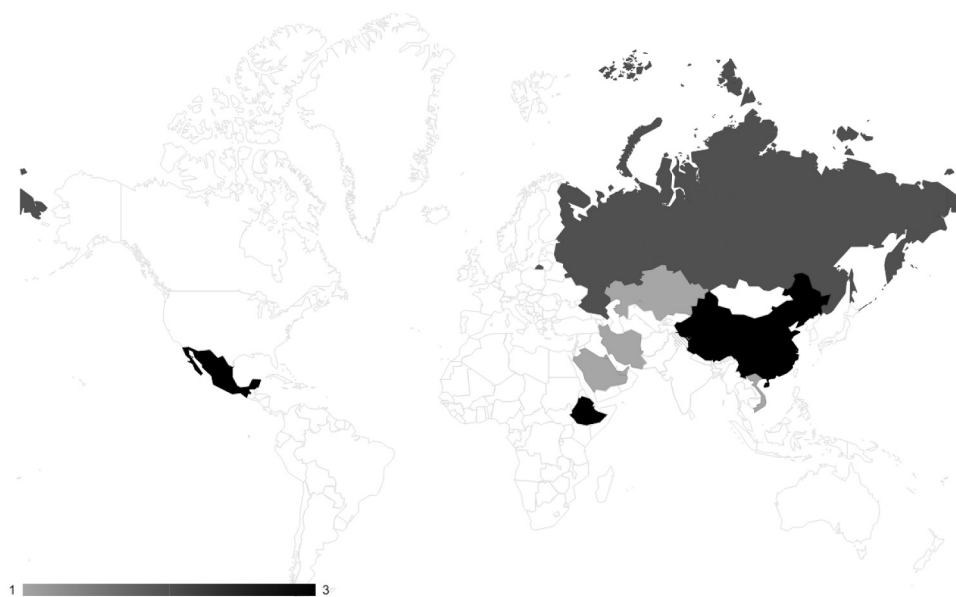
**Figure 7.** Commercial reporting – attribution of targeted threats to civil society.

‘cyber powers’, hence it is conceivable that they are the main perpetrators of threats to civil society.

Accordingly, independent reporting (Figure 8) also covers six campaigns by these ‘big three’, underlining their importance. However, it also documents the use of targeted digital threats by a range of other governments absent from commercial reporting: Kazakhstan (1), Ethiopia (3), Kuwait (1), Saudi Arabia (1), United Arab Emirates (1) and Bahrain (1),

and Mexico (3). Independent reporting shows not only a more evenly distributed attribution pattern, but the total number of operations by ‘other’ states is actually greater than those attributed to the ‘big three’. These findings strongly support the hypothesized bias toward adversarial actors.

Our findings thus provide strong support for hypothesis 1, 2 and 3; the theory of bias in commercial reporting passes the hoop test.



**Figure 8.** Independent reporting – attribution of targeted threats to civil society.

### Case studies

This section develops a plausibility probe of our theory. Our case studies test the hypothesized selection bias in commercial reporting in favor of unique TTP, high-profile victim, and high-profile perpetrators against evidence from three case studies of cyber operations targeting civil society.

#### Case 1: *Tainted Leaks*

The *Tainted Leaks* operation targeted a journalist and involved the theft of personal e-mails and subsequent ‘leaking’ of this data (Hulcoop, Scott-Railton, Tanchak, Brooks, & Deibert, 2017). It is a most-likely case with high values across all three selection criteria, and reporting patterns are congruent with expectations: commercial reporting prominently covers the threat actor, but not its targeting of civil society. Moreover, even reports that do discuss civil society focus on high-profile targets while bracketing lower-profile targeting. In this case, we trace such omissions in detail, impossible in the other two cases due to a lack of available commercial reporting on the respective campaigns/threat actors.

This operation is a most-likely case with extreme values across all three selection criteria. First, it involved unique TTP of ‘tainting’ leaked data, by carefully including disinformation within otherwise legitimate data. Second, it involves high-profile targets. The victim, journalist David Satter, is a high-profile Kremlin

critic, but a low-profile actor concerning the sector’s revenue potential. The operation was part of a larger scale phishing operation against several high-profile targets, including “a former Russian Prime Minister, members of cabinets from Europe and Eurasia, ambassadors, high ranking military officers, CEOs of energy companies” (Hulcoop et al., 2017). Third, the operation was pursued by a high-profile actor. Citizen Lab found circumstantial evidence pointing to APT28, and based on “additional evidence” Forbes confirmed this suspected attribution one day later (Fox-Brewster, 2017). APT28 has been attributed to the Russian military intelligence agency GRU (US vs. NETYKSHO et al, 2018), it has received extensive media coverage – especially since its intrusion into the DNC (CrowdStrike, 2016) – and is consistently ranked among the most dangerous threat actors (Burgess, 2017; NCSC, 2018). Due to the high-profile actor, high-profile victims and unique TTP, we would expect a high volume of reporting on this actor and the campaign the *Tainted Leaks* operation was part of.

Reporting volumes are congruent with expectations: a significant proportion of threat reports discuss APT28 operations, 57 out of the 630 reports in our dataset (9%). Considering there are over forty known threat actors, this disproportionate attention to APT28 supports the hypothesized reporting bias toward high-profile threat actors.

Prioritization of civil society targeting is also in line with predictions. Unfortunately, few commercial threat

reports include targeting proportions. However, the Citizen Lab investigation did include such proportions: in the phishing campaign associated with the Tainted Leaks case, civil society was the second-largest target group at 21%, behind only governments, comprising 24% of targets (Hulcoop et al., 2017). Those few commercial reports with targeting proportions confirm this prioritization of civil society. SecureWorks' investigation of an APT28 phishing campaign in 2016 using similar methods found that most targets (41%) were military, but the second largest target group were civil society actors at 36% (including NGOs, activists, and journalists) (SecureWorks, 2016b). Moreover, a 2015 TrendMicro report identified civil society as the main target of APT28's domestic operations (TrendMicro, 2015b). Finally, the first commercial report on APT28, published by FireEye in 2014, explicitly highlighted its targeting of journalists to "monitor public opinion, identify dissidents, spread disinformation or facilitate further targeting." (FireEye, 2014). Evidently, civil society is a priority target of APT28, hence unbiased commercial reporting patterns should reflect this prioritization.

Out of 57 commercial reports on APT28, however, only 15 mention civil society (26%), and only two out of these prioritize a threat to civil society. The great majority of reporting (39 reports, 74% of all reporting on APT28) entirely omits threats to civil society. If one were to build an analysis of APT28 activity entirely based on commercial reporting, civil society would appear a low priority, or only an occasional target of this actor.

Meanwhile, congruent with our predictions, reporting prioritizes high-profile targets. Most reports characterize APT28 as a highly sophisticated espionage actor targeting governments and the private sector. For example, CrowdStrike describes it as a "Russian-based threat actor . . . responsible for targeted intrusion campaigns against the Aerospace, Defense, Energy, Government and Media sectors." (CrowdStrike, 2016). Similarly, Symantec states that "the organizations targeted by APT28 during 2017 and 2018 include: a well-known international organization, military targets in Europe, Governments in Europe, a government of a South American country, an embassy belonging to an Eastern European country" (Symantec, 2018). Notably, this periodization includes the *Tainted Leaks* campaign, yet neither the latter, nor the wider spear phishing campaign against civil society it is part of are

mentioned by Symantec. In short, the prioritization of government and military targets in commercial reporting is congruent with our predicted selection bias toward high-profile targets. The apparent omission of civil society targeting in most reporting provides further support for our theory. Recall that its targeting of dissidents was mentioned in very first report on APT28, and available data on vertical targeting patterns suggest civil society is among its main targets. It would thus be highly surprising if this pattern is entirely absent from a majority of its reported activity.

Moreover, prioritization of high-profile targets is evident not only in the reports that do not mention civil society, but also in those that do. For example, a FireEye survey of APT28 activity in 2014–2017 only mentions one civil society target: the dissident band Pussy Riot (FireEye, 2017, p. 4), which has received widespread media coverage. A 2015 TrendMicro also leads with the targeting of Pussy Riot, and reveals the APT28's broad targeting of civil society (TrendMicro, 2015b). Yet even this report is scant on the details, instead prioritizing high-profile victims in its analysis, noting that "to illustrate one of the credential phishing attacks Pawn Storm [APT28] sends to its targets, we will focus on a particular attack on high-profile Yahoo users" (TrendMicro, 2015b).

Finally, only a small proportion of commercial reporting mentions the credential phishing operation preceding the *Tainted Leaks* campaign. This finding is congruent with the expected selection bias toward unique TTP. Most reports on APT28 focus on sophisticated methods, few mention the relatively simple deception involved in the phishing campaign. Nine commercial reports mention credential phishing, seven of which do discuss civil society. Two out of the latter discuss the specific technique of credential phishing used in the Tainted Leaks operation and against Hillary Clinton's campaign (SecureWorks, 2016b, 2016a). Importantly, SecureWorks only published its first report one day after CrowdStrike had revealed APT28's breach of the DNC, which provided a high-profile target and corresponding media attention. The timing of publication provides additional support for the hypothesized prioritization of high-profile targets.

Meanwhile, no commercial reports mention the technique of 'tainting leaks'. One report briefly highlights 'alleged' data manipulation, noting how "prior to leaking the information [obtained by APT28], parts of the documents and e-mails were allegedly altered. "

(TrendMicro, 2017, p. 6). Yet it provides no further details about the data involved, the method or possible aims.

In short, reporting patterns are largely congruent with expectations. Commercial reporting frames APT28 as a highly sophisticated espionage actor targeting governments and large private sector entities. There are references to civil society targeting, but only among a minority of reports. The targeting of journalists to spread disinformation was indicated from the beginning of reporting on APT28 in 2014, hence one would expect operations such as *Tainted Leaks* to be reported prominently. Yet commercial reporting mostly brackets civil society, and does not discuss ‘tainted leaking’ apart from a passing mention. Findings in this most-likely case confirm the hypothesized selection bias.

### **Case 2: Spying on a Budget**

Our second case study lies at the opposite end of the spectrum concerning selection criteria. The *Spying on a Budget* report by Citizen Lab identifies a spear phishing campaign targeting the Tibetan community of activists. It involves a low-profile actor using generic methods to go after (mostly) low-profile targets (Crete-Nishihata, Dalek, Maynier, & Scott-Railton, 2018). These properties make it a most-likely case, where our theory predicts a low volume or complete lack of reporting on the actor in question and/or this specific campaign. Findings confirm this expectation, no commercial report mentions either the actor or the campaign – hence precluding an analysis of prioritization of civil society targeting within commercial reporting.

*Spying on a Budget* analyzes a phishing operation active from around January 2016 until July 2017 using a range of tactics to obtain the e-mail credentials of members from the Tibetan activist community and potentially other social movements in China (Crete-Nishihata et al., 2018). These individual activists and civil society groups constitute low-profile targets promising little business opportunities and media attention. However, while the decoy documents employed indicate the Tibetan community as the main target, additional documents used indicate the same campaign also targeted government agencies in South and Southeast Asia. The presence of higher-profile victims increases the likelihood of threat reporting. The operation’s methods were cheap and unsophisticated; the

Citizen Lab estimates a total budget of only 1000 USD. Finally, the threat actor involved is unknown, exhibits “only basic technical skills” and is sloppy, leading Citizen Lab researchers to conclude it is likely a ‘low-level contractor’ (Crete-Nishihata et al., 2018). In short, it is a very low-profile threat actor, highly unlikely to fulfil the third selection criterion for publication in threat reporting. Hence, our theory would predict only few, if any, passing mentions of this campaign and/or actor in commercial reporting.

Results are congruent with our expectations. There are no preceding commercial reports on this threat actor and/or the phishing campaign involved. Moreover, there are also no follow-up reports after publication of the Citizen Lab report. These findings are in line with our expectations. However, RecordedFuture published a thoroughly researched report on a campaign targeting the Tibetan community six months after Citizen Lab’s report, attributing it to the same threat actor (Recorded Future, 2018). This level of attention to the same low-level actor would challenge our theory, were it not for the fact of the overall “increased level of sophistication for the attacker” (Recorded Future, 2018). Evidently, the threat actor had passed the necessary threshold in uniqueness and sophistication to be included in a threat report. Findings on the overall lack of commercial reporting on this threat actor, as well as the inclusion only following an increase in sophistication, are congruent with our theory’s predictions.

### **Case 3: Familiar Feeling**

The *Familiar Feeling* operation also targets Tibetan activists, but is pursued by a somewhat higher-profile actor that also targets additional, somewhat higher-profile victims, and employs somewhat unique TTP. Because it involves a medium-profile actor using TTP of medium sophistication going after medium-profile targets, this is a least-likely case where our theory does not provide strong predictions either way. Findings concerning reporting volume are inconclusive: there are five commercial reports on this threat actor, less than on high-profile actors like APT28, but more than the lowest profile actors – such as in the case above. Findings on the prioritization of civil society targeting within these five reports, however, strongly confirm the expected selection bias: none of them mention civil society – even though one of the reports specifically

mentions a key piece of evidence indicating civil society targeting.

The *Familiar Feeling* campaign involves a known threat actor, known as Tropic Trooper or KeyBoy (Alexander et al., 2018), which was also behind a 2016 campaign targeting the Tibetan community (Hulcoop, Brooks, Maynier, Scott-Railton, & Crete-Nishihata, 2016). Tropic Trooper has not been conclusively linked to a specific government, but is suspected to be associated with China. It has not received attention in general news media. However, dedicated information security media has covered previous activity by this actor (Muncaster, 2017; Networks Asia Staff, 2015). Its campaign reported by Citizen Lab – dubbed ‘Resurfaced’ – employed a new version of a previously known set of malware exploiting known vulnerabilities. Finally, while this campaign focuses on low-profile victims (an unnamed Tibetan NGO), Citizen Lab links it to preceding TropicTrooper/KeyBoy campaigns against government and large private sector actors in East and Southeast Asia. Hence, it involves both low and high(er)-profile targets.

In short, this campaign falls in the middle of the spectrum of selection criteria, with a medium level of sophistication, a medium-profile threat actor and a low-profile target but previous campaigns against higher profile targets. Hence, it is a least-likely case where our theory only weakly predicts reporting outcomes: it is possible some commercial reporting covers this campaign since it may cross the necessary thresholds for publication, but the opposite outcome is similarly likely. However, since this campaign includes both lower and higher-profile victims, any evidence for a prioritization of high-profile targets – and in particular omissions of lower profile targets – provides strong support for our theory.

Threat reporting patterns on this case clearly support our expectations. Overall, five commercial reports discuss campaigns by TropicTrooper, which by itself does not confirm or challenge our expectations. Significantly, however, none of them mention civil society targeting, instead exclusively focusing on high-profile government and corporate targets. Rapid7 first reported on the actor in 2013, vaguely hypothesizing targeting of “either someone in the telecommunications industry or a representative of the local government” (Rapid7, 2013). TrendMicro reported targeting of “major government sectors and corporations in both

Taiwan and the Philippines” (TrendMicro, 2015a). PwC’s report quotes previous Citizen Lab research on the actor but without mentioning civil society. Instead, while noting the lack of “clear visibility” into targeting, it nonetheless highlights that it “does appear that this latest campaign targets at least some Western organizations, likely for corporate espionage purposes” (PwC, 2017) – providing support for the hypothesized prioritization of victims in the Global North. Based on commercial reporting, one would thus conclude that this is an actor focusing exclusively on international espionage.

There are three plausible explanations for this exclusive focus on high-profile government and corporate actors: (1) the campaigns reported by Citizen Lab are the only ones targeting civil society, (2) commercial researchers were unaware of the targeting of civil society, or (3), commercial researchers were aware but did not include it in reporting due to the prioritization of high-profile victims. In the former two cases, we would not expect any evidence pointing toward civil society targeting commercial threat reporting. However, there are two key pieces of such evidence – and congruent with the third explanation. First, the reference to Citizen Lab research in the PwC report shows its authors were familiar with Tropic Trooper’s targeting of civil society. Second, the latest report characterizes it as an actor “focusing on . . . government, healthcare, transportation, and high-tech industries” and reports on the evolution of its tradecraft (TrendMicro, 2018). However, that report’s ‘indicators of compromise’ section also includes the domain “tibet-news[.]today”, pointing directly toward the targeting of Tibetan community by the same actor (later shown in the Resurfaced campaign by Citizen Lab). Yet the TrendMicro report does not address this piece of evidence and its implications. To be sure, none of these findings provide conclusive ‘smoking gun’ evidence of selection bias in favor of high-profile targets in commercial reporting. However, both overall reporting patterns and anecdotal pieces of evidence pointing to the omission of civil society targeting are closely congruent with our predictions and thus strongly support our theory.

In conclusion, none of the findings in the most-likely cases challenge our theory, while findings in our least-likely case strongly confirm our expectations. Our hypotheses pass the plausibility probe.



## Discussion

Both the quantitative and qualitative analysis support our hypotheses about threat reporting sample bias. Overall reporting patterns, as well as the cases examined, are congruent with predictions based on our theory of threefold selection bias. While the limitations of available data prevent a causal analysis, the unambiguousness of our findings – in particular the least-likely case – strongly indicate that reporting prioritizes sophisticated and unique campaigns by high-profile threat actors against high-profile targets. Conversely, the cybersecurity marketplace fails to provide sufficient reporting at the low-end of cyber conflict. Cybersecurity firms are guilty of failures of omission rather than commission – firms are focusing preferentially on particular classes of threats rather than actively discriminating against another. This situation has two important implications.

First, commercial reporting creates a distorted picture of cyber conflict as researchers base their analyzes on a skewed sample of cases. There is growing evidence that cyber conflict thrives especially at the low end of the conflict spectrum (Lindsay, 2017), and in this conflict civil society is right at the frontlines (R. Deibert, 2015). Yet, our findings suggest this portion of conflict is systematically sidelined in threat reporting.

Second, this distorted picture poses a risk for democracy by systematically underrepresenting the threats to the CSOs that are vital for the functioning of democracy. Indeed, it seems increasingly likely that the original cyberwar narrative had things precisely backwards. The information revolution does not portend a new anarchy rife with destructive disruption but rather the encroaching hierarchy of the surveillance state. Cyberspace may create asymmetric advantages, but they are advantages of the strong to monitor and enforce the behavior of the weak. The good news about a lower likelihood of cyberwar is expressly bad news for democratic liberties and human rights.

Moreover, this distorted picture implies a linear relationship between technical sophistication and threat level that does not hold in practice. Gioe et al. argue that cybersecurity firms typically focus on the technical aspects of security “because they are relatively easier to secure”, although most cyber operations exploit weaknesses in human cognition and “do not need high-end nation state cyber tools to

achieve their goals” (Gioe, Goodman, & Wanless, 2019, p. 118). These are precisely the types of threats that we have shown to be underreported, underlining the need to transition from a technical, state-centric conception of cybersecurity toward a human-centric approach (Deibert, 2018).

The solution to this problem cannot come from the market alone, yet governments are simultaneously key threats to civil society. Several threat intelligence firms are offering pro-bono services to civil society, which is a move to be welcomed, but these individual measures cannot override the market logic that dictates the priorities of the commercial security sector as a whole. We have already alluded to the public goods nature of public threat reporting at the outset of this article.<sup>19</sup> Public goods theory tells us that if such goods are provided by private actors, they will be insufficiently and unevenly distributed because benefits are shared among the group while costs are borne by the individual actor(s) providing it alone (Olson, 1971, p. 34–36). As Olson shows, these characteristics lead to a classic collective action problem: a market failure manifested in the overrepresentation of concentrated interests and underprovision of diffuse benefits. The classic solution to such cases where markets fail to allocate resources efficiently and evenly is state intervention (Hardin, 2015, p. 52). Yet in this case, state security agencies are the main threats to civil society, and CSOs require independence from governments. Consequently, a government-driven solution is not a viable option. After all, government-sponsored support to CSOs abroad challenging authoritarian regimes constitutes a form of interference that those regimes can perceive as a cyber-attack on their vital interests in domestic political stability.

Conversely, widespread surprise at the methods used in Russia’s meddling in the 2016 U.S. election attests to a fundamental lack of awareness of the vulnerabilities of democratic institutions and civil society more broadly that are emerging in the deepening information revolution. Russia did not invent civil society surveillance, suppression, and disruption, and authoritarian actors will continue to find new ways to leverage cyberspace. The distorted understanding of the nature of cyber threats has resulted in (1) lower than necessary prioritization and resource-allocation for cyber defense in public policy, as well as (2) insufficient preparedness by

both policy-makers and civil society itself when it comes to detecting and mitigating these threats.

Our findings highlight the need for a follow-up analysis with statistical methods on these and additional data as well as ethnographic engagement with cybersecurity firms to gauge significance of the selection criteria identified here. In particular, the impact of the level of sophistication as well as the profile of the threat actor on reporting volume needs to be analyzed more systematically. This task faces two key challenges: first, establishing a general measure of sophistication, including both technical and social aspects, and second, it requires consolidating the naming schemes to track reporting. Currently, each firm employs their own naming schemes, and there are no commonly accepted criteria for sophistication.

The best available solution to close the information gap is awareness of the limitations of commercial research, as well as increased independent research of targeted threats across the entire spectrum of cyber conflict. There is an urgent need for more interdisciplinary research into targeted threats with academic rigor and transparency of methods and selection criteria. This analysis points to the need for foundations and funders that are often the principal supporters of civil society to take notice of these targeted digital threats and take measures to mitigate them through their grant-making. There are signs of change, such as the Ford Foundation's digital security initiative (Brennan, Eagen, Nunez, Scott-Railton, & Sears, 2017), but we are still far from a broad recognition and prioritization of this issue.

## Notes

1. See the next section for more details.
2. For more details and a definition of these threats, please see Online Appendix, Section A1.
3. A definition is included in the Online Appendix, Section A1.
4. For examples of this influence, please see Online Appendix, Section A2.1.
5. These include Citizen Lab, Electronic Frontier Foundation, AccessNow, Human Rights Watch and Amnesty International.
6. See Online Appendix, Section A1, for definitions of these terms.
7. This dual role also allows for a less cynical interpretation of threat reporting as a quasi-academic enterprise, yet with similar results. See Online Appendix, Section A2.2 for more details.
8. There are two such projects: (1) APT Groups and Operations, a sheet consolidating naming schemes and operations; and (2) APTNotes, a repository of commercial reporting.
9. See Online Appendix, Section A2.3 for more details.
10. See Online Appendix, Section A2.4 for a more detailed discussion of these expectations.
11. See Online Appendix, Section A3.3, for further details on the assumed causal mechanism and limitations in the availability of data.
12. See Online Appendix, Section A3.1, for further details on these criteria.
13. A coding guide is available in the Online Appendix, Section A3.4, which also discusses reliability measures and provides a link to our data.
14. See Online Appendix, Section A3.6, for further details.
15. Our qualitative analysis tracks these additional indicators as well.
16. For coding details on attribution, see Online Appendix, Section A3.4.
17. See Online Appendix, Section A3.2 for a discussion of rival theories.
18. Some commercial reports discuss spyware in general (Cf. Kaspersky 2017), but not its targeted use.
19. Public goods are defined by two key properties: they are non-exclusive (Samuelson, 1954), and non-rivalrous (Ostrom & Ostrom, 1977). In other words, no one can be excluded from the benefits of the good, while its consumption by one actor does not reduce the availability to others. Public threat reporting fulfills both criteria: it is freely available online and reading a report does not reduce availability to others (Rosenzweig, 2011).

## Acknowledgments

The authors would like to thank Max Smeets, Masashi Crete-Nishita, Irene Poetranto, Adam Casey, and Alexei Abrahams for their insightful comments on earlier drafts of this paper. We also thank the participants of the 2018 "Global Digital Futures" workshop at Columbia University's School of International and Public Affairs, the team at ETH Zurich's Center for Security Studies, participants of the 2019 ISA panel on Digital Technologies and Human Rights, and the Ostrom workshop at Indiana University Bloomington for their helpful feedback. Daria Goriacheva provided excellent research assistance for the reliability test. We are grateful for the generous funding from the Carnegie Corporation of New York and the School of International and Public Affairs at Columbia University, the Ford Foundation, the John D. and Catherine T. MacArthur Foundation, the Sigrid Rausing Trust, the Oak Foundation and the Open Society Foundations that helped make this project possible. Finally, we thank AccessNow for providing us with aggregate data, and in particular Daniel Bedoya for his help in preparing this data.

## Declaration of interest statement

The authors declare there are no conflicts of interest.

## Funding

This work was supported by the Carnegie Corporation of New York; Ford Foundation; John D. and Catherine T. MacArthur Foundation; Oak Foundation; Open Society Foundations; Sigrid Rausing Trust.

## ORCID

Lennart Maschmeyer  <http://orcid.org/0000-0003-4666-2387>

## Notes on contributors

*Lennart Maschmeyer* is a Senior Researcher at the Center for Security Studies, ETH Zurich.

*Ronald J. Deibert* is a Professor of Political Science, and Director of the Citizen Lab at the Munk School of Global Affairs & Public Policy, University of Toronto.

*Jon R. Lindsay* is an Assistant Professor of Digital Media and Global Affairs at the Munk School of Global Affairs and Public Policy, University of Toronto.

## Data availability statement

A copy of the dataset can be accessed here: [https://docs.google.com/spreadsheets/d/1FyzBsZ1Uvhr2inK\\_cKItBgSzY7JlGHDLbrEQKFD08w/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1FyzBsZ1Uvhr2inK_cKItBgSzY7JlGHDLbrEQKFD08w/edit?usp=sharing)

## References

- Alexander, G., Brooks, M., Crete-Nishihata, M., Maynier, E., Scott-Railton, J., & Deibert, R. J. (2018). *Familiar feeling: A Malware campaign targeting the tibetan diaspora resurfaces*. Retrieved from <https://citizenlab.ca/2018/08/familiar-feeling-a-malware-campaign-targeting-the-tibetan-diaspora-resurfaces/>
- Anderson, R., Barton, C., Böhme, R., Clayton, R., van Eeten, M. J. G., Levi, M., ... Savage, S. (2013). Measuring the cost of cybercrime. In B. Rainer. (Ed.), *The economics of information security and privacy* (pp. 265-300). Berlin Heidelberg, Germany: Springer-Verlag. Retrieved from <https://www.springer.com/gp/book/9783642394973>
- Arquilla, J., & Ronfeldt, D. (1993). Cyberwar is coming! *Comparative Strategy*, 12(2), 141–165. doi:10.1080/01495939308402915
- Bator, F. M. (1958). The anatomy of market failure. *The Quarterly Journal of Economics*, 72(3), 351–379. doi:10.2307/1882231

- Borghard, E. D., & Lonergan, S. W. (2017). The logic of coercion in cyberspace. *Security Studies*, 26(3), 452–481. doi:10.1080/09636412.2017.1306396
- Brantly, A. F. (2014). The cyber losers. *Democracy and Security*, 10(2), 132–155. doi:10.1080/17419166.2014.890520
- Brennan, M., Eagen, E., Nunez, B., Scott-Railton, J., & Sears, E. (2017). *Digital security and grantcraft guide*. Retrieved from <https://www.fordfoundation.org/media/3334/digital-security-grantcraft-guide-v10-final-22317.pdf>
- Burgess, M. (2017). *Exposed: How one of Russia's most sophisticated hacking groups operates*. *Wired UK*. Retrieved from <https://www.wired.co.uk/article/how-russian-hackers-work>
- CLTC. (2018). Center for long-term cybersecurity project on protecting politically vulnerable organizations threat landscape and organizational ecosystem. UC Berkeley.
- Crete-Nishihata, Jakob Dalek, Ronald Deibert, Seth Hardy, Katherine Kleemola, Sarah McKune, Irene Poetranto, John Scott-Railton, Adam Senft, Byron Sonne, and Greg Wiseman. 'Communities @Risk: Targeted Digital Threats Against Civil Society'. Citizen Lab, 11 November 2014. <https://targetedthreats.net/>. <https://targetedthreats.net/media/1-ExecutiveSummary.pdf>.
- Crete-Nishihata, M., Dalek, J., Maynier, E., & Scott-Railton, J. (2018). *Spying on a budget: Inside a phishing operation with targets in the tibetan community*. The Citizen Lab. Retrieved from <https://citizenlab.ca/2018/01/spying-on-a-budget-inside-a-phishing-operation-with-targets-in-the-tibetan-community/>
- CrowdStrike. (2016). Bears in the Midst: Intrusion into the democratic national committee »." Retrieved from <https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee>
- CrowdStrike. 2019. Global threat report. Retrieved from <https://www.crowdstrike.com/blog/2019-global-threat-report-shows-it-takes-innovation-and-speed-to-win-against-adversaries/>
- Deibert, R. (2015). Cyberspace under siege. *Journal of Democracy*, 26(3), 64–78. doi:10.1353/jod.2015.0051
- Deibert, R., & Rohozinski, R. (2009). *Tracking GhostNet: Investigating a cyber espionage network*. *The Citizen Lab* (blog). Retrieved from <https://citizenlab.org/2009/03/track-ing-ghostnet-investigating-a-cyber-espionage-network/>
- Deibert, R. J. (2003). Black code: Censorship, surveillance, and the militarisation of cyberspace. *Millennium - Journal of International Studies*, 32(3), 501–530. doi:10.1177/03058298030320030801
- Deibert, Ronald J. 'Toward a Human-Centric Approach to Cybersecurity'. *Ethics & International Affairs* 32, no. 4 (ed 2018): 411–24. <https://doi.org/10.1017/S0892679418000618>.
- Dunn Cavelt, M. (2013). From cyber-bombs to political fall-out: Threat representations with an impact in the cyber-security discourse. *International Studies Review*, 15(1), 105–122. doi:10.1111/misr.12023
- Dunn-Cavelt, M. (2008). Cyber-terror—looming threat or phantom menace? The framing of the US cyber-threat debate. *Journal of Information Technology & Politics*, 4(1), 19–36. doi:10.1300/J516v04n01\_03

- Egloff, F. J. (2020). Contested public attributions of cyber incidents and the role of academia. *Contemporary Security Policy*, 41(1), 55–81. doi:10.1080/13523260.2019.1677324
- ESET. (2010). Stuxnet under the microscope.
- FireEye. (2014). *APT28: A window into Russia's cyber espionage operations?* FireEye. Retrieved from <https://www2.fireeye.com/rs/fireeye/images/rpt-apt28.pdf>
- FireEye. (2017). *APT28: At the center of the storm*. FireEye. Retrieved from <https://www2.fireeye.com/rs/848-DID-242/images/APT28-Center-of-Storm-2017.pdf>
- FireEye. (2019). *M-Trends*. Retrieved from <https://content.fireeye.com/m-trends/rpt-m-trends-2019>
- Fox-Brewster, T. (2017). *Russian 'Fancy Bear' hackers tainted their huge leaks with fake data*. *Forbes*. Retrieved from <https://www.forbes.com/sites/thomasbrewster/2017/05/26/russian-dnc-hackers-planted-leaks-with-fake-data/>
- Future, R. (2018). *RedAlpha: New campaigns discovered targeting the tibetan community*. Retrieved from <https://www.recordedfuture.com/redalpha-cyber-campaigns/>
- Gartzke, E. (2013). The myth of cyberwar: Bringing war in cyberspace back down to earth. *International Security*, 38(2), 41–73. doi:10.1162/ISEC\_a\_00136
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, Mass: MIT Press.
- Gioe, D. V., Goodman, M. S., & Wanless, A. (2019). Rebalancing cybersecurity imperatives: patching the social layer. *Journal of Cyber Policy*, 4(1), 117–137. doi:10.1080/23738871.2019.1604780
- Guerrero-Saade, J. A. (2015). *The ethics and perils of APT research*. Retrieved from <https://media.kaspersky.com/pdf/Guerrero-Saade-VB2015.pdf>
- Hardin, R. (2015). *Collective action*. RFF Press. doi:10.4324/9781315044330.
- HBGary. (2010). *Operation Aurora*.
- Hulcoop, A., Brooks, M., Maynier, E., Scott-Railton, J., & Crete-Nishihata, M. 2016. *It's parliamentary: KeyBoy and the targeting of the tibetan community*. Retrieved from <https://citizenlab.ca/2016/11/parliament-keyboy/>
- Hulcoop, A., Scott-Railton, J., Tanchak, P., Brooks, M., & Deibert, R. 2017. *Tainted leaks: Disinformation and phishing with a Russian nexus*. Retrieved from <https://citizenlab.ca/2017/05/tainted-leaks-disinformation-phish/>
- Isaac, M., & Wakabayashi, D. 2017. Russian influence reached 126 million through facebook alone. *The New York Times*. sec. Technology. Retrieved from <https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html>
- Jayamaha, B., & Matissek, F. (2018). *Hybrid war: Attacking the 'Civil' in civil society*. *US Army War College War Room* (blog). Retrieved from <https://warroom.armywarcollege.edu/articles/hybrid-war-attacking-the-civil-in-civil-society/>
- Jensen, B., Valeriano, B., & Maness, R. (2019). Fancy bears and digital trolls: Cyber strategy with a Russian twist. *Journal of Strategic Studies*, 42(2), 212–234. doi:10.1080/01402390.2018.1559152
- Kaspersky. (2017). *Use of commercial android spyware almost doubled in 2017*. Retrieved from [https://www.kaspersky.com/about/press-releases/2017\\_use-of-commercial-android-spyware-almost-doubled-in-2017](https://www.kaspersky.com/about/press-releases/2017_use-of-commercial-android-spyware-almost-doubled-in-2017)
- Kuerbis, B., & Badiei, F. (2017). Mapping the cybersecurity institutional landscape. *Digital Policy, Regulation and Governance*, 19(6), 466–492. doi:10.1108/DPRG-05-2017-0024
- Langner, R. (2011). What stuxnet is all about. *Langner - Reliable and Safe IIOT (Industrial Internet of Things)* (blog). Retrieved from <https://www.langner.com/2011/01/what-stuxnet-is-all-about/>
- Lawson, S. (2013). Beyond cyber-doom: Assessing the limits of hypothetical scenarios in the framing of cyber-threats. *Journal of Information Technology & Politics*, 10(1), 86–103. doi:10.1080/19331681.2012.759059
- Lieberman, E. S. (2005). Nested analysis as a mixed-method strategy for comparative research. *American Political Science Review*, 99(3), 435–452. doi:10.1017/S0003055405051762
- Lindsay, J. R. (2013). Stuxnet and the limits of cyber warfare. *Security Studies*, 22(3), 365–404. doi:10.1080/09636412.2013.816122
- Lindsay, J. R. (2014). The impact of China on cybersecurity: Fiction and friction. *International Security*, 39(3), 7–47. doi:10.1162/ISEC\_a\_00189
- Lindsay, J. R. (2017). Restrained by design: The political economy of cybersecurity. *Digital Policy, Regulation and Governance*, 19(6), 493–514. doi:10.1108/DPRG-05-2017-0023
- Lynn, W. J., III (2010). Defending a New Domain. *Foreign Affairs*. Retrieved from <http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain>
- Marczak, B., Scott-Railton, J., Senft, A., Poetranto, I., & Sarah, M. (2015). Pay no attention to the server behind the proxy - Mapping FinFisher's continuing proliferation. *The Citizen Lab* (blog). Retrieved from <https://citizenlab.org/2015/10/mapping-finfishers-continuing-proliferation/>
- McAfee. (2010). *Operation Aurora - post mortem*. McAfee Blogs. Retrieved from <https://securingtomorrow.mcafee.com/business/security-connected/operation-aurora-%e2%80%93post-mortem/>
- Mueller, M., Grindal, K., Kuerbis, B., & Badiei, F. (2019, April). Cyber attribution: Can a new institution achieve transnational credibility? *Cyber Defense Review*, (pp. 107–121).
- Mueller, R. S., III. (2019). *Report on the investigation into Russian interference in the 2016 presidential election*. Retrieved from <https://www.justice.gov/storage/report.pdf>
- Mukherjee, S. P. (2018). *Statistical methods in social science research*. New York, NY: Springer Berlin Heidelberg.
- Muncaster, P. (2017). *Chinese KeyBoy group unlocks more victim networks*. *Infosecurity Magazine*. Retrieved from <https://www.infosecurity-magazine.com/443/news/chinese-keyboy-group-unlocks/>
- NCSC. (2018). *Indicators of compromise for malware used by APT28 - NCSC site*. Retrieved from <https://www.ncsc.gov.uk/alerts/indicators-compromise-malware-used-apt28>
- Networks Asia Staff. (2015). *Philippines, Taiwan are latest targets of 'operation tropic trooper' Malware*. *Networks Asia*. Retrieved from <https://www.networksasia.net/article/philippines-taiwan-are-latest-targets-operation-tropic-trooper-malware.1432083840>
- Nye, J. S. (2011). *The future of power*. 1st ed. New York, USA: PublicAffairs.

- ODNI. (2017). *ODNI statement on declassified intelligence . . . IC on the record*. Retrieved from <https://icontherecord.tumblr.com/post/155494946443/odni-statement-on-declassified-intelligence>
- Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups* (Revised ed., pp. 124). Cambridge, Massachusetts; London, England: Harvard University Press, Harvard Economic Studies.
- Ostrom, V., & Ostrom, E. (1977). *Public goods and public choices*. Indiana, USA: Indiana University, Workshop in Political Theory and Policy Analysis.
- PwC. (2017). *The KeyBoys are back in town*. Retrieved from <https://www.pwc.co.uk/issues/cyber-security-data-privacy/research/the-keyboys-are-back-in-town.html>
- Rapid7. (2013). *KeyBoy, targeted attacks against Vietnam and India*. Retrieved from <https://blog.rapid7.com/2013/06/07/keyboy-targeted-attacks-against-vietnam-and-india/>
- Rattray, G. J. (2009). An environmental approach to understanding cyberpower. In *Cyberpower and national security* Franklin D. Kramer, Stuart H. Starr, Larry K. Wentz (Eds.), (pp. 253–274). Washington, D.C.: Potomac Books.
- Rid, T. (2012). Cyber war will not take place. *Journal of Strategic Studies*, 35(1), 5–32. doi:10.1080/01402390.2011.608939
- Rid, T., & Buchanan, B. (2015). Attributing cyber attacks. *Journal of Strategic Studies*, 38(1–2), 4–37. doi:10.1080/01402390.2014.977382
- Rogin, Josh. ‘Opinion | Washington Must Wake up to the Abuse of Software That Kills’. *Washington Post*, 2018. <https://www.washingtonpost.com/opinions/2018/12/12/washington-must-wake-up-abuse-software-that-kills/>.
- Rosenzweig, P. (2011). *Cybersecurity and public goods*. Retrieved from [http://media.hoover.org/sites/default/files/documents/EmergingThreats\\_Rosenzweig.pdf](http://media.hoover.org/sites/default/files/documents/EmergingThreats_Rosenzweig.pdf)
- Samuelson, P. A. (1954). The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4), 387–389. doi:10.2307/1925895
- Schleifer, T., & Walsh, D. (2017). *McCain: Russian cyberintrusions an ‘act of war*. CNN. Retrieved from <http://www.cnn.com/2016/12/30/politics/mccain-cyber-hearing/index.html>
- SecureWorks. (2016a). *Hillary clinton email targeted by threat group-4127*. Retrieved from <https://www.secureworks.com/research/threat-group-4127-targets-hillary-clinton-presidential-campaign>
- SecureWorks. (2016b). *Threat group-4127 targets google accounts*. Retrieved from <https://www.secureworks.com/research/threat-group-4127-targets-google-accounts>
- Shezaf, Hagar. ‘Snowden: Israeli Firm’s Spyware Was Used to Track Khashoggi - Israel News - Haaretz.Com’, 2018. <https://www.haaretz.com/israel-news/premium-israeli-spyware-was-used-to-track-saudi-journalist-khashoggi-edward-snowden-says-1.6633745>.
- Slayton, R. (2017). What is the cyber offense-defense balance? Conceptions, causes, and assessment. *International Security*, 41(3), 72–109. doi:10.1162/ISEC\_a\_00267
- Smeets, M. (2018). A matter of time: On the transitory nature of cyberweapons. *Journal of Strategic Studies*, 41(1–2), 6–32. doi:10.1080/01402390.2017.1288107
- Symantec. (2011). W32.Stuxnet Dossier.
- Symantec. (2018). APT28: New espionage operations target military and government organizations. Retrieved from <https://www.symantec.com/blogs/election-security/apt28-espionage-military-government>
- Symantec. (2019). *Internet security threat report*. Retrieved from <https://www.symantec.com/content/dam/symantec/docs/reports/istr-24-executive-summary-en.pdf>
- Threat Intelligence Researcher. 2018Phone.
- TrendMicro. (2015a). Operation tropic trooper: Old Vulnerabilities Still Pack a Punch - TrendLabs Security Intelligence Blog. Retrieved from <http://blog.trendmicro.com/trendlabs-security-intelligence/operation-tropic-trooper-old-vulnerabilities-still-pack-a-punch/>
- TrendMicro. (2015b). Pawn storm’s domestic spying campaign revealed; Ukraine and US top global targets - trendlabs security intelligence blog. Retrieved from <https://blog.trendmicro.com/trendlabs-security-intelligence/pawn-storms-domestic-spying-campaign-revealed-ukraine-and-us-top-global-targets/>
- TrendMicro. (2017). *Two years of pawn storm*. Retrieved from <https://documents.trendmicro.com/assets/wp/wp-two-years-of-pawn-storm.pdf>
- TrendMicro. (2018). *Tropic trooper’s new strategy - trendlabs security intelligence blog*. Retrieved from <https://blog.trendmicro.com/trendlabs-security-intelligence/tropic-trooper-new-strategy/>
- US vs. NETYKSHO et al. (2018). US District Court of District of Columbia.
- Van Evera, S. (1997). *Guide to methods for students of political science*. Ithaca, NY: Cornell University Press.
- Work, J. D. (2020). Evaluating commercial cyber intelligence activity. *International Journal of Intelligence and CounterIntelligence*, 1–31. doi:10.1080/08850607.2019.1690877
- YouGov. (2017). *America’s friends and enemies | YouGov*. Retrieved from <https://today.yougov.com/topics/politics/articles-reports/2017/02/02/americas-friends-and-enemies>