The Journal of Strategic Studies, 2015
Vol. 38, Nos. 1–2, 4–37, http://dx.doi.org/10.1080/01402390.2014.977382

Routledge
Taylor & Francis Group

Attributing Cyber Attacks

THOMAS RID AND BEN BUCHANAN

Department of War Studies, King's College London, UK

ABSTRACT Who did it? Attribution is fundamental. Human lives and the security of the state may depend on ascribing agency to an agent. In the context of computer network intrusions, attribution is commonly seen as one of the most intractable technical problems, as either solvable or not solvable, and as dependent mainly on the available forensic evidence. But is it? Is this a productive understanding of attribution? — This article argues that attribution is what states make of it. To show how, we introduce the Q Model: designed to explain, guide, and improve the making of attribution. Matching an offender to an offence is an exercise in minimising uncertainty on three levels: tactically, attribution is an art as well as a science; operationally, attribution is a nuanced process not a black-and-white problem; and strategically, attribution is a function of what is at stake politically. Successful attribution requires a range of skills on all levels, careful management, time, leadership, stress-testing, prudent communication, and recognising limitations and challenges.

KEY WORDS: Cyber Security, Attribution, Traceability, Information Security, Signals Intelligence

Attribution is the art of answering a question as old as crime and punishment: who did it? Doing attribution well is at the core of virtually all forms of coercion and deterrence, international and domestic. Doing it poorly undermines a state's credibility, its effectiveness, and ultimately its liberty and its security.

Decisions of life and death depend on attribution. The use of chemical weapons in Ghouta, a suburb of Damascus, in August 2013; the downing of Malaysia Airlines Flight 17 near Donetsk Oblast, Ukraine, in the summer of 2014; the abduction of three Israeli teenagers in Gush Etzion in June, which triggered the Gaza War of 2014 — all these events have in common that nobody immediately claimed credit, and that the identity of the perpetrators remained highly contested *while* consequential political decisions had to be made at the highest levels. The attribution problem has not raised its profile so dramatically only in recent years. The assassination of Archduke Franz Ferdinand of Austria on 28 June 1914 offered a similar conundrum: who was Gavrilo Princip, the assassin? And was he an agent of the Serbian state?

Attribution unwinds incrementally. These international incidents illustrate the potentially enormous stakes at play. But they are too exceptional and too confusing for a systematic discussion of attribution. Beginning with a more orderly and established illustration is more productive. In law enforcement, identifying a felon may begin with a report of a crime to an emergency phone operator. Next come investigators. The officers will secure the scene and interview witnesses. Forensic specialists will try to find and analyse specific artefacts, for instance matching a bullet found in the victim to a gun with fingerprints found at the crime scene. If all goes well, the evidence will be marshalled into a case presented to a jury, where the final question of attribution will be settled. Though often fraught with drama, it is a methodical, ordered, and institutionalised approach.

This scenario is simplistic but instructive. It reveals at least three general and familiar features: attribution is almost always too large and too complex for any single person to handle; attribution is likely to require a division of labour, with specialities and sub-specialities throughout; and attribution proceeds incrementally on different levels, immediate technical collection of evidence, follow-up investigations and analysis, and then legal proceedings and making a case against competing evidence in front of a decision authority. The law enforcement scenario is extensively explored in scholarly literature and popular culture. Attributing cyber attacks is less simple and the ground less familiar.

In cyber security, the attribution debate is evolving surprisingly slowly.[1] Three common assumptions currently dominate the discussion on digital attribution. The first assumption is that attribution is one of the most intractable problems[2] of an emerging field, created by the underlying technical architecture[3] and geography[4] of the Internet.

[1]For an early contribution, see, David A. Wheeler and Gregory N. Larsen, *Techniques for Cyber Attack Attribution* (Alexandria, VA: Institute for Defense Analysis 2003); Richard Clayton, *Anonymity and Traceability in Cyberspace*, vol. 653, *Technical Report* (Cambridge: Univ. of Cambridge Computer Laboratory 2005); Susan Brenner, At Light Speed": Attribution and Response to Cybercrime/Terrorism/Warfare', *The Journal of Criminal Law & Criminology*. 97/2 (2007), 379–475. For an early case study, see, Clifford Stoll, *The Cuckoo's Egg* (New York: Doubleday 1989).

[2]'Perhaps the most difficult problem is that of attribution', P.W. Singer and Allan Friedman, *Cybersecurity and Cyberwar* (New York/Oxford: OUP Press, 2014, p. 73. See also, David Betz and Tim Stevens, *Cyberspace and the State*, Adelphi Series (London: IISS/Routledge 2011), 75–6.

[3]See for instance, W. Earl Boebert, 'A Survey of Challenges in Attribution', in Committee on Deterring Cyberattacks (ed.), *Proceedings of a Workshop on Deterring Cyberattacks* (Washington DC: National Academies Press 2011), 51–2. Also, Martin Libicki, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND Corporation 2009), 43.

[4]Jack Goldsmith and Tim Wu, *Who Controls the Internet? Illusions of a Borderless World* (Oxford: OUP 2006).

Only a technical redesign of the Internet, consequently, could fully fix the problem.[5] Similar positions prevail in the legal debate.[6] The second assumption is a binary view on attribution: for any given case, the problem can either be solved,[7] or not be solved.[8] Either attribution leads to the culprit, or at some point it simply ends with a spoofed IP address, obfuscated log files, or some other dead trail.[9] The third common assumption is that the attributive evidence is readily comprehensible, that the main challenge is finding the evidence itself, not analysing, enriching, and presenting it.[10] These views are common; they are intuitive; and they are not wrong — but they are limited and insufficient. The reality of attribution has evolved significantly in the past decade. Actual attribution of cyber events is already more nuanced, more common, and more political than the literature has acknowledged so far.[11]

---

[5]Mike McConnell, 'How to Win the Cyberwar We're Losing', *Washington Post*, 28 Feb. 2010.

[6]See, Matthew C. Waxman, 'Cyber-Attacks and the Use of Force', *The Yale Journal of International Law* 36 (2011),. 421–59, 447; Nicholas Tsagourias, 'Cyber Attacks, Self-Defence and the Problem of Attribution', *Journal of Conflict & Security Law* 17 (2013), 229–44. For a discusson on levels of attribution necessary for the use of force, see Marco Roscini, *Cyber Operations and the Use of Force in International Law* (Oxford: OUP 2014), 33–40.

[7]Former Secretary of Defense Leon Panetta famously said on the *USS Intrepid*, 'the [DoD] has made significant advances in solving a problem that makes deterring cyber adversaries more complex: the difficulty of identifying the origins of an attack.' Leon Panetta, Remarks on Cybersecurity to the Business Executives for National Security, New York City', Washington DC: Department of Defense, 12 Oct. 2012.

[8]David D. Clark and Susan Landau, 'Untangling Attribution', in Committee on Deterring Cyberattacks (ed.), *Proceedings of a Workshop on Deterring Cyberattacks*, (Washington DC: National Academies Press 2011). See also Jason Healey, *A Fierce Domain* (Washington DC: The Atlantic Council 2013), 265.

[9]Robert K. Knake, 'Untangling Attribution: Moving to Accountability in Cyberspace, Planning for the Future of Cyber Attack', Washington DC: Subcommittee on Technology and Innovation, 111th Congress, 15 July 2010.

[10]The most influential articles on intrusion analysis seem to assume that the evidence speaks for itself, as they do not focus on the problem of communicating results to a non-technical audience. The two most influential and useful contributions are the 'Diamond Model', see Sergio Caltagirone, Andrew Pendergast and Christopher Betz, *The Diamond Model of Intrusion Analysis,* ADA586960 (Hanover, MD: Center for Cyber Threat Intelligence and Threat Research 5 July 2013), as well as the 'Kill Chain' analysis, see, Eric M. Hutchins, Michael J. Cloppert and Rohan M. Amin, *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains* (Bethesda, MD: Lockheed Martin Corporation 2010).

[11]See Boebert, 'A Survey of Challenges in Attribution', 41–54. For a wider perspective, see, Amir Lupovici, 'The "Attribution Problem" and the Social Construction of "Violence"', *International Studies Perspectives* 2014, 1–21.

This article attempts to move the debate on attribution beyond these entrenched positions. It raises three sets of questions. We start by considering the first-order question: if attribution is not first and foremost a technical problem, what is it instead? A second question follows accordingly: if attribution is not a binary affair but a matter of degree, what, then, is normal attribution and how is high-quality attribution different from low-quality attribution? And third: if evidence is inconspicuous and equivocal, how should it be marshalled and analysed? How should attribution as a whole be managed and communicated to third parties?

This text argues that *attribution is what states make of it*. Matching an offender to an offence is an exercise in minimising uncertainty on several levels. On a technical level, *attribution is an art as much as a science*. There is no one recipe for correct attribution, no one methodology or flow-chart or check-list. Finding the right clues requires a disciplined focus on a set of detailed questions — but also the intuition of technically experienced operators. It requires *coup d'œil*, to use a well-established military term of art.[12] On an operational level, *attribution is a nuanced process, not a simple problem*. That process of attribution is not binary, but measured in uneven degrees, it is not black-and-white, yes-or-no, but appears in shades. As a result, it is also a team sport — successful attribution requires more skills and resources than any single mind can offer. Optimising outcomes requires careful management and organisational process. On a strategic level, *attribution is a function of what is at stake politically*. The political stakes are determined by a range of factors, most importantly by the incurred damage. That damage can be financial, physical, or reputational. Viewed from the top, attribution is part resourcing and guiding the internal process; part participating in final assessments and decisions; and part communicating the outcome to third parties and the public.

To grasp the argument and illustrate an idealised making of attribution, we introduce the Q Model (see Figure 1).[13] Tactically, the model helps analysts ask the full range of relevant questions, to aid their critical thinking, and to put an investigation into context.[14] Operationally, the model helps integrate both technical and non-

[12]Carl von Clausewitz used *coup d'œil* to describe 'military genius,' the 'inward eye' that enables good commanders to make the right tactical decisions under stress, information overload, and time-constraints, see, Carl von Clausewitz, *On War*, translated by Michael Howard and Peter Paret (Princeton UP 1976), 100–12.

[13]Q alludes to a number of things: first and foremost it hints at questions, the crux of attribution. Q also links to quartermaster, a type of naval officer with particular responsibility for signals and steering. The etymological root of 'cyber' is κυβερνώ (kyvernǫ́), *to steer*.

[14]The model is deliberately designed neither as a flowchart nor as a checklist. In several focus group sessions with operators it became clear that any linear representation would not be able to reflect the uniqueness and varied flow of the wide range of cases investigators handle.

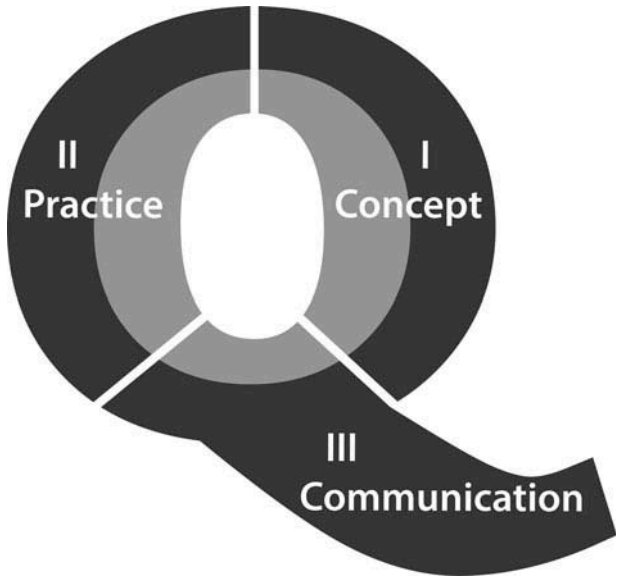

Figure 1. Structure of this article and of the detailed graph (see annex).

technical information into competing hypotheses. This includes asking more challenging questions on different levels, including fine-grained, detail-driven technical questions as well as broader, more analytical operational questions. Strategically, the model helps refine and extract the essence of the attribution process for external presentation in the appropriate estimative language. This language may inform political judgements with serious consequences.

Figure 1 illustrates how this article will proceed, how the argument will be presented, and how to read the model's far more detailed graphic illustration provided in the annex. The first part is conceptual: it will introduce attribution as a process by discussing the model in general terms and introducing several critical distinctions and dynamics. The second part is empirical: it will illustrate various steps along the attribution process through recent examples. The third part will consider the proverbial hook that protrudes from the Q's base, the challenge of communicating the potentials and limitations of attribution and translating the findings into action. The conclusion takes stock, reassesses several entrenched yet problematic views, and considers the limitations of attributing cyber attacks.

Part I

This study is designed as a conceptual and practical map for mastering the attribution process on all levels, from forensic investigators to

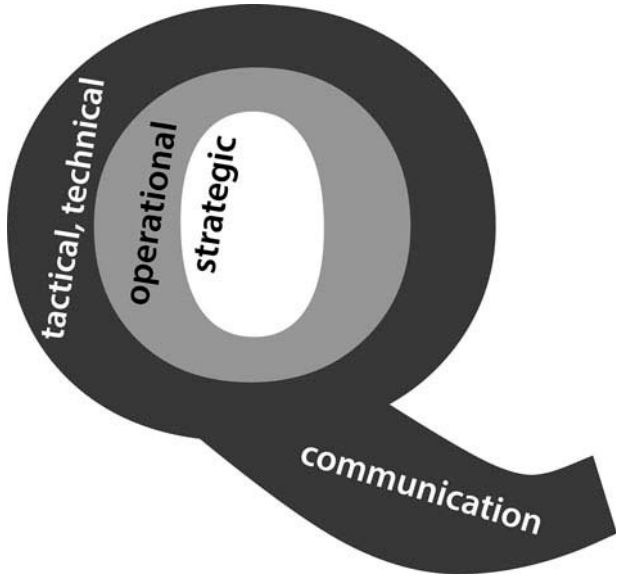

Figure 2. The three layers of analysis.

intelligence analysts to national security officials, executives, political leaders, to journalists and scholars writing about cyber security and network intrusions.

Each of the levels of the attribution process represents a discrete analytical challenge, relies on specific input data and specific expertise, and illuminates a separate aspect of successful attribution (see Figure 2). The analysis on each level needs to be informed and tempered by the others. Though the attribution process typically has a beginning and an end, the cycle does not necessarily follow a specific sequential or chronological order, as hypotheses are confronted with new details and new details give rise to new hypotheses in turn. Nevertheless, the layers represent separate tasks that, though they interrelate, will be analysed individually here. Usually so-called 'indicators of compromise' trigger the attribution process. Such indicators raise specific technical questions. More questions are likely to follow only after more facts have been gathered. On occasion, the attribution process may begin on the operational or strategic level. Sometimes the 'first knowers' of an incident will be above the technical level. Guided by non-forensic sources of intelligence, or by the broader geopolitical context — sometimes even by intuition — the possibility of malicious activity may be identified before technical indicators flag it, or indeed even before it begins. Attribution can go either way: the strategic and operational layers may inform the subsequent technical analysis, or vice versa.

Broad skills are required for attribution. Cyber threats have reached a high level of complexity. Both executing them and uncovering their architecture through attributive analysis requires a refined division of labour. A team of anti-virus researchers, for instance, can spend considerable time and energy reverse-engineering the installation mechanism of a specific piece of malware, while control engineers may focus on the particular design of a target-specific payload against an industrial plant. Stuxnet was so complex that particular companies focused their analysis on different aspects, such as the propagation mechanism, the command-and-control setup, or the payload targeting the industrial control system.[15] As in a military context, an entire range of tactical activities lies beneath — but is vital to — operational considerations. Analysing the separate aspects requires vastly different skills — this specialisation is a firmly established principle in criminal investigations as well as in military operations: no commander would put IED-disposal units in charge of analysing the financing of insurgent networks, or the supply-chain of explosive devices. F-16 pilots do not choose their own targets. Missile engineers do not do nuclear strategy. In the context of cyber attacks, such elementary expectations have yet to form outside the small expert community engaged with attribution work.

The overall goals of the attribution process often depend on the incurred damage, or potential damage. In a world of many incidents and not enough investigators, the amount of damage caused or threatened frequently determines the resources that are invested into attributing the malicious event itself. If an intrusion did not cause any obvious damage, a company or even a government agency may decide to ignore it, to only partially investigate it, or perhaps to improve its defences generally but not launch an expensive investigation into the origins of the seemingly inconsequential breach. A lack of perceived damage can thus short-circuit the attribution process before it even fully starts. To some degree, this is unavoidable.

The tactical goal is understanding the incident primarily in its technical aspects, the *how*. The operational goal is understanding the attack's high-level architecture and the attacker's profile — the *what*. The strategic goal is understanding who is responsible for the attack, assessing the attack's rationale, significance, appropriate response — the *who* and *why*. Finally *communication* is also a goal on its own: communicating the outcome of a labour-intensive forensic investigation is part and parcel of the attribution process, and should not be treated as low priority. Indeed public attribution itself can have significant

[15]For an overview see Jon R. Lindsay, 'Stuxnet and the Limits of Cyber Warfare', *Security Studies* 22/3 (2013), 365–404.

effects: offenders may abort an operation, change tactics, or react publicly to allegations, thus shaping the victim's wider response.

Detail is critical. But detail can also overwhelm. As information flows from the technical to the operational and strategic layers, it must be synthesised. Only then will it be comprehensible and useful. Technical analyses can, depending on the incident, yield a prodigious amount of detail about specific intrusions. This will often include the specific exploits that were used, the payload mechanism, the command-and-control infrastructure, the targeted data, reverse-engineering analysis, and raw data from the affected networks. Some, maybe even most, of the technical details collected will have limited relevance. Some tactically relevant details may lose their significance on operational and strategic levels, just as details of geopolitical context are of limited concern to the forensic investigator. This process extracts meaning from the detail: absent proper synthesis, a high density of technical forensic artefacts does not necessarily mean that operational or strategic questions can be answered with more certainty. Detail is not fungible within the larger attribution process.

Certainty, therefore, is uneven. As the information flows from technical to operational to strategic, the questions get sparser and broader. Thus the uncertainty of attributive statements is likely to increase as the analysis moves from technical to political. *What was the intrusion mechanism?* is a question that can be answered based on forensic artefacts. *What was the motive?* is a query that will require developing hypotheses, and the subsequent testing of such hypotheses.[16] A technical forensic question may be narrowly focused and concretely answerable. Competing operational hypotheses may be informed by labour-intensive forensic evaluations, but not fully backed by the available technical and non-technical evidence. On a strategic level conclusions are yet further removed from forensic artefacts, and may contain a significant amount of assumptions and judgement.[17] Educating senior decision-makers is vital to managing this problem.

Aperture comes in here. One of the most difficult elements in the attribution process is what many in the intelligence community call *aperture*: the scope of sources that can be brought to bear on a specific

[16]The exception may be some forms of crime. Identifying a monetary incentive is easier than examining a political incentive.

[17]Staff with a more abstract and formal training, for instance those with a mathematical background, may be inclined to formalise cyber security problems. This can be counterproductive. Abstraction can conceal a lack of insight. For an example of highly questionable formalisation and faux-precision, see Robert Axelrod and Rumen Iliev, 'Timing of cyber conflict', *PNAS* 111/4 (28 Jan. 2014), 1298–303. Even the mathematical formalism in one widely used model for intrusion analysis, the so-called 'Diamond Model,' may imply an exaggerated degree of precision. Caltagirone, Pendergast and Betz, *The Diamond Model of Intrusion Analysis*.

investigation, akin to the variable opening through which light enters a camera. The quality of attribution is likely to rise as the number of fruitful intelligence sources increases. Moreover, the significance of a wider aperture rises with the levels of the attribution process: opening the aperture on a specific incident on a purely technical level is possible, but only within narrow constraints. Digital forensic evidence generated by an intrusion is by definition limited in the context it provides. Exploit code rarely reveals motivation. On an operational and especially on a strategic level, other sources of intelligence may illuminate the wider picture, for instance intercepted telephone conversations or emails among those who ordered or organised an operation. The significance of all-source intelligence and of a wider aperture is one of the strongest reasons why states with highly capable intelligence agencies are better equipped to master the attribution process than even highly capable private entities.

The very first large-scale state-on-state computer network intrusion set in history, MOONLIGHT MAZE, demonstrates the value of all-source intelligence and a wide aperture. The intrusions came to light in 1998.[18] Foreign spies targeted the US Department of Defense (DoD), Department of Energy, National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), various defence contractors, and universities. The intruders exfiltrated information ranging from helmet designs to atmospheric data. FBI investigators initially were overwhelmed. In early 1999, the DoD began supporting the investigation. The intelligence directorate in the Joint Task Force Computer Network Defense, JTF-CND, 'left no stone unturned' — they started with the digital forensic data obtained by law enforcement investigators, but then included signals intelligence, human intelligence, even the history of overhead imagery of specific suspected buildings to see if they recently had communications equipment installed. Ultimately intelligence sources that went beyond the digital forensic artefacts of the actual intrusions enabled attributing the MOONLIGHT MAZE breaches to the Russian government with a reasonable level of certainty.[19]

Individual persons can gain significance when attributing network breaches. If evidence can be produced that links an intrusion to an individual within an organisation, then the attribution will be stronger. This contrasts starkly with many international incidents, especially military incidents: many weapon systems and capabilities are marked, soldiers wear uniforms, and often the geography of an incident points to

[18]For an overview of MOONLIGHT MAZE, see Adam Elkus's chapter in Healey, *A Fierce Domain*, 152–63.

[19]Author interviews with former members of JTF-CND and the FBI's MOONLIGHT MAZE Task Force, Washington DC, Sept to Nov, 2014.

the identity of the organisation behind an intrusion. A specific military target, for example an experimental nuclear facility in Syria, may get hit from the air. Syria, or even third parties, could identify the raiding F-15s as part of the Israeli Air Force through geographical context or aircraft type or flight paths — all without identifying the individual pilots. Military organisations can be identified without identifying individuals and smaller units first — this may be starkly different in cyber operations.

The ultimate goal of attribution is identifying an organisation or government, not individuals. But in lieu of markings, uniforms, and geography, individual operators can be powerful links between malicious artefacts and organisations. One of the most useful examples is CrowdStrike's Putter Panda report, published on 9 June 2014. One of the company's UK-based researchers, Nathaniel Hartley, first identified a malicious actor who was using the handle 'cpyy' in coordinated breaches. The next step was linking 'cpyy' to a real person. Hartley used registration data to connect the handle to Chen Ping. Now Chen, or 'cpyy', had to be connected to an organisation. Hartley uncovered more identifying information from various sources, including blogs and a Picasa photo album. Pictures in the folder 'office' clearly linked Chen aka 'cpyy' to a building in Shanghai, through various details in the images, including military hats, buildings, equipment, and even portraits of Chen. With the help of these photos, Hartley pinpointed a location: 31°17′17.02′N longitude 121°27′14.51′E, in the heart of the Zhabei District of Shanghai. The address represented the headquarters of the People's Liberation Army's (PLA) General Staff Department, 3rd Department, 12th Bureau, Unit 61486.[20] Hartley's evidence combines multiple sources and is convincing.[21] Other examples are Mandiant's APT1 report and, in a more limited sense, an exceptional Department of Justice indictment.[22] All of these reports construct links between individuals and organisations via their online personas. On its own, such a personal link may not be sufficient for high-quality attribution. Yet credibly identifying an organisation may require first zooming down to persona level — and then zooming back out to organisational or unit level. This dynamic will depend on the available aperture. If the personal link aligns with other indicators from other sources, then the evidence can strengthen the case significantly.

[20]Nathaniel Hartley, *Hat-tribution to PLA Unit 61486,* CrowdStrike, 9 June 2014; see also *Putter Panda,* CrowdStrike, 9 June 2014.

[21]Author communication, by email, 6 Aug. 2014. The significance of persona research is highly controversial among the leading cyber security firms, with FireEye and Kaspersky being more sceptical. Focus group session with FireEye staff, Reston, VA, 15 Sept. 2014 and with Kaspersky staff, Barcelona, 8 Oct. 2014.

[22]The indictment will be discussed in some detail later in this paper.

Perception also matters. Preconceptions, prejudgments, prejudice, and psychological and political biases are likely to influence attribution. This dynamic has an internal and an external aspect: internally, analysts and managers at all levels may be inclined to produce the expected findings and interpret evidence in a specific light. Organisational dynamics can amplify this problem as internal reports are passed up. 'Policy premises constrict preception, and administrative workloads constrain reflection', as one prominent study of intelligence failures found in 1978.[23] A synthetic example may illustrate this point: the Saudi government could hypothetically discover that the 'Cutting Sword of Justice,' a group that credibly claimed the 2012 Shamoon attack against Saudi Aramco, consists of a small number of Saudi-based Shia activists. Possibly prejudiced against Shia activists in a Sunni majority country, Saudi investigators could be tempted to assume that the group was tasked by authorities in Iran, even if the available evidence would not fully support linking Saudi citizens of Shia background to Tehran. The bigger the internal perception bias, the bigger is the risk of costly mistakes.

Part II

The quality of attribution is a function of asking the right questions. Each of the model's layers has its specific set of queries that drive the process on that level. The answers to questions from one layer inform the starting points on the next. The better a team's overview of the entire process, the better is the quality of the attribution. This process is dynamic and non-linear: each case is different, so any rigid flow-model or linear 'checklist' approach to an investigation is problematic.[24] The following paragraphs will discuss the process layer by layer, starting with tactical-technical considerations and slowly moving up to strategic considerations. If possible, each aspect will be illustrated with very short references to empirical examples.[25]

The technical layer is often the starting point of an investigation. It is both broad and deep. This places great challenges on staff. Analysts are expected to work in an efficient, team-oriented manner to answer questions about computer code, network activity, language, and much more.

[23]Richard K. Betts, 'Analysis, War, and Decision: Why Intelligence Failures Are Inevitable', *World Politics*, 31/1 (Oct. 1978), 61–89, 61.

[24]Analysts repeatedly and unanimously voiced scepticism towards linear 'checklists' in a number of focus group sessions in the private and public sectors over the summer of 2014.

[25]We will not have space to introduce these examples in detail, and will therefore provide references to the most authoritative source in each case. These sources are sometimes academic publications, but more often company reports.

The technical evidence in many cases forms the basis of the attribution process. Unearthing this evidence is not always glamorous, but vital.

Indicators of compromise are likely to begin an investigation. Indicators of compromise are technical artefacts of network intrusion or malicious activity, often abbreviated as IOCs in technical jargon. Such indicators are typically uncovered either through broad-based automated scanning or reports of aberrant computer behaviour. Performing deep, individualised, and regular forensic analysis on a large number of computers is often too costly for network administrators. IOCs serve as a useful heuristic to help narrow the scope of follow-up investigations. One influential study divided the indicators of compromise into three main categories: atomic, behavioural, and computed.[26]

Atomic indicators are discrete pieces of data that cannot be broken down into their components without losing their forensic value. Atomic indicators, by themselves, pinpoint malicious activity. Common ones include IP addresses, email addresses, domain names, and small pieces of text. Computed indicators are similarly discrete pieces of data, but they involve some element of computation. An example is a 'hash', a unique signature derived from input data, for instance a password or a program. A hash is always the same value as long as the input does not change. Hashes of programs running on their network's computers may match hashes of programs known to be malicious. Behavioural indicators are combinations of actions and other indicators that both reveal malicious activity and — in some cases — point to a specific adversary who has employed similar behaviours in the past. A behavioural indicator might be repeated social engineering attempts of a specific style via email against low-level employees to gain a foothold in the network, followed by unauthorised remote desktop connections to other computers on the network delivering specific malware. Organisations that are careful about computer network defence collect all three types of indicators of compromise and routinely scan their network and computers for them. Once evidence of a compromise is found, more technical questions follow. Their order will vary based on the indicator, adversary, and threat.

Almost all intruders must overcome one challenge: entry.[27] Any attacker must acquire the ability to execute code on an unauthorised system. Such code will exploit a system vulnerability and grant the attacker further access or functionality. A common way to deliver this code does not exploit technical vulnerabilities, but human weakness:

[26]Eric M. Hutchins, Michael J. Cloppert and Rohan M. Amin, *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains* (Bethesda, MD: Lockheed Martin Corporation 2010), 3.

[27]An exception is denial of service attacks. These seek to deny availability of certain computer systems by overwhelming them with basic, often meaningless, data.

spear-phishing, the practice of sending socially-engineered messages in order to trick the user into taking some action. A famous breach of the security firm RSA began with an email sent to small groups of low-level employees. The email, entitled '2011 Recruitment Plan' was convincing enough that one of the employees retrieved it from the junk mail folder and opened the attachment. It was an Excel file containing a malicious exploit, permitting the attackers access to the system. From this beachhead, they moved through the network to more valuable computers.[28] Such occurrences are fairly common, even against high profile targets,[29] and investigators look to them to see what clues to attribution they might provide. Technical data are associated with spear-phishing, such as the origin of the email, but so are social data, such as language mistakes and sophistication in targeting. Another entry method relies on USB drives infected with remote access software. These can either be inserted by the attacker or an associated agent, or by an unwitting employee of the target using a spiked USB device. There are more purely technical methods of entry as well. A common approach is a watering-hole attack. This approach requires hacking a web site likely to be visited by the target — something as benign as the site of a takeout restaurant[30] — so that when the targeted employee visits the site, his or her computer is breached via a vulnerability in the web browser. A number of entry techniques rely on manipulating and compromising legitimate web requests to a benign site by controlling network infrastructure, either as a so-called 'man-in-the-middle attack' or, if the attacker does not control a node but can still inject data, as a 'man-on-the-side' attack.[31]

Targeting can shed light on the type of breach or the type of intruder. Credit card information and other easily monetised targets point to organised criminals. Product designs may point to a range of competing companies in countries engaged in economic espionage. Details on

[28]Uri Rivner, *Anatomy of an Attack,* RSA, 1 April 2011.

[29]According to an internal State Department cable made public by WikiLeaks, 'Since 2002, [US government] organizations have been targeted with social-engineering online attacks' which resulted in 'gaining access to hundreds of [US government] and cleared defense contractor systems'. Brian Grow, and Mark Hosenball, 'Special report: In cyberspy vs. cyberspy, China has the edge', *Reuters*, 14 April 2011.

[30]Nicole Perlroth, 'Hackers Lurking in Vents and Soda Machines', *New York Times*, 8 April 2014, A1.

[31]For an example, 'Is This MITM Attack to Gmail's SSL?', Google Product Forums, <http://bitly.com/alibo-mitm+>; also Seth Schoen and Eva Galperin, 'Iranian Man-in-the-Middle Attack Against Google Demonstrates Dangerous Weakness of Certificate Authorities', *Electronic Frontier Foundation*, 29 Aug. 2011. See also Nicholas Weaver, 'A Close Look at the NSA's Most Powerful Internet Attack Tool', *Wired*, 13 March 2014.

political and military strategy can point to intelligence agencies. The technical layer can provide specific artefacts related to targeting that will inform working assumptions on an operational layer. By looking at an intruder's movement between computers in a breached network, for instance, investigators may gain insight into what the attackers were after. By reconstructing specific commands issued by the attacker, investigators may be able to see from the memory of infected machines if the attackers had something specific in mind, or if they were looking broadly for anything that might be of value. Sometimes code also contains search-terms: one operation known as Ouroboros, revealed in 2014, contained the search terms 'NATO' and 'EU Energy Dialogue'.[32]

Targeting analysis can also help illuminate the organisational setup of the attacker. The resources the attacker brought to bear in the effort may be an indicator for how highly the attacker valued the target. If an attack uses many more resources than it needs to — for example, a sophisticated rootkit for low-level espionage — this can be a sign that the operation is less likely to be part of a group that values efficiency in its targeting. A similar indicator can be redundant targeting: some attackers may use the same methodology on the same target multiple times, even after one breach attempt has already succeeded. Such redundancy of effort may be an indicator that the attacker represents a large organisation, possibly with a confused tasking setup. This 'spraying' of large numbers of targets may also indicate a division of labour between breachers and exploiters on the part of the attacker.[33]

Infrastructure is required for most malicious activities. In the case of a denial of service attack, which relies on overwhelming the targeted computer with meaningless information, the infrastructure actually performs the attack. In other cases of malicious activity, infrastructure is often used as a jumping-off point or to issue instructions to code on compromised machines (command-and-control in technical jargon). To maximise efficiency and minimise logistical costs, malicious entities will often reuse this physical digital infrastructure from one breach to another. It therefore can be a valuable clue in the attribution process, establishing links between different operations and potentially between different groups. In the American indictment of five PLA officers, prosecutors specifically cited the operators' usage of domain name servers as part of their attribution process.[34] An offender can acquire various

[32] *The Epic Turla Operation,* Kaspersky Lab, 7 Aug. 2014.
[33] Author interviews with various operators, Summer 2014.
[34] *United States of America v Wang Dong, Sun Kailiang, Wen Xinyu, Huang Zhenyu, Gu Chunhui,* Criminal Nr 14-118, Erie, PA: US District Court Western District of Pennsylvania, 1 May 2014, Exhibit F.

degrees of infrastructure ownership: a computer, for example, could be hijacked as a 'bot' without its legitimate owner taking notice; a server could be rented legitimately from a service provider and then used for malicious purposes; or infrastructure could be owned and physically maintained by the attacker. The type of link that an attacker has to the enabling infrastructure determines follow-on questions in a number of ways. For example, rented infrastructure, such as virtual machines and servers, may open up access to more registration and log information through service providers. Infrastructure that an adversary owns and maintains could lead to clues about the adversary's physical location. In any case, monitoring an adversary's infrastructure can open new trails of analysis and help interdict future operations. As a result, some shrewd actors are taking steps to try to better hide their infrastructure.[35]

Modularity is one of the most prominent features of computer code. For reasons of efficiency, malicious actors will often avoid reinventing the wheel and will re-use software to accomplish basic tasks in their operations. As part of an attack, this software is frequently loaded directly on to the target networks, where it can later be analysed by investigators. Often, his software has its own signatures and hallmarks, which can provide insight into the identity of the intruders and their supporters. FireEye, a leading security company, illustrates as much with their report on so-called Digital Quartermasters. These are enabling entities that provide the same software to a range of affiliated malicious groups.[36] This sort of analysis can vary in utility, depending on the case. Some code, packaged in modules, is so commonly used in intrusions that it ceases to be a very useful indicator for identifying an offender. Other code, like the underlying code for both the Stuxnet and Duqu malware, is so esoteric or complex as to be very useful in identification.[37] In those investigations, researchers were reasonably certain that the authors of Stuxnet also authored Duqu, because the two pieces of malware shared some key modules and the code was not widely available.[38]

[35]On 6 Aug. 2014, for instance, FireEye disclosed an operation in which 'malware appears to beacon to legitimate domains', in an attempt to 'lull defenders into a false sense of security', see Ned Moran, Joshua Homan and Mike Scott, *Operation Poisoned Hurricane,* FireEye, 6 Aug. 2014.

[36]Ned Moran and James Bennett, *Supply Chain Analysis: From Quartermaster to Sunshop,* FireEye Labs, 11 Nov. 2013.

[37]See Costin Raiu, 'Inside the Duqu Command and Control Servers', presentation at SOURCE Boston 2012, 4 May 2012, <http://youtu.be/nWB_5KC7LE0>.

[38]The Symantec report on Duqu notes, 'Duqu shares a great deal of code with Stuxnet; however, the payload is completely different. Instead of a payload designed to sabotage an industrial control system, it has been replaced with general remote access capabilities. The creators of Duqu had access to the source code of Stuxnet, not just the Stuxnet binaries [compiled versions],' *W32.Duqu,* Version 1.4, Symantec, 23 Nov. 2011, 3.

The pattern-of-life of intrusions is an important part of breach investigations. All organisations rely on schedules and routines in order to maximise efficiency. Hacking groups are no exception. Timing and other patterns of activity can thus give clues to their location and identity. For example, the United States indictment against five PLA members indicates that the operators followed a reasonably set schedule of work. The operatives set their command-and-control infrastructure to only ping for instructions during those hours, going so far as to turn them off during lunchtime, overnight, and on weekends.[39] This behaviour matched the business hours in Shanghai, where the US government alleges the intruders are located. In another example, CrowdStrike attributed an offensive campaign to attackers in Russia, because most of the compilation times — the moment when software is packaged for use — occurred during working hours in Russia.[40] Patterns-of-life are easy to fake, yet widely used to corroborate working assumptions of investigators.

Language indicators in malware can also provide clues for attribution. There are two main categories of language artefacts: those that reveal words chosen by an attacker for a specific thing — such as names of variables, folders, and files — and computer artefacts revealing general configuration settings. Either is relatively easy to fake in a sophisticated 'false flag' operation. Yet language analysis nonetheless remains a worthy part of the attribution process. Examples abound, but a recent one is the Careto malware discovered by Kaspersky. 'Careto' was the name given by the malware authors to one of two main modules of their espionage vehicle. As is common, the operation's command-and-control servers were scattered across a large number of countries, the majority of them in non-Spanish-speaking countries like Malaysia, Austria, the Czech Republic, Singapore, and the United States. But the language artefacts told a different story.[41] The first indicator was a number of subdomains that purported to be Spanish newspapers, probably used for spear-phishing (though British and American newspapers were also impersonated).[42] A second indicator was that the configuration data revealed the code was developed on machines with Spanish language settings. A third indicator was slang words that, the Russian researchers suspected, 'would be very uncommon

[39] *United States of America v Wang Dong, Sun Kailiang, Wen Xinyu, Huang Zhenyu, Gu Chunhui,* Criminal Nr 14-118, Erie, PA: US District Court Western District of Pennsylvania, 1 May 2014, 12–13, Exhibit F.

[40] Author interview with Dmitri Alperovich, Arlington, VA, 15 Sept. 2014, see also *Global Threat Report,* Arlington, VA: CrowdStrike, 22 Jan. 2014, 18.

[41] *Unveiling 'Careto',* Version 1.0, Kaspersky Lab, 6 Feb. 2014, 46.

[42] For example, elpais.linkconf[dot]net/ and elespectador.linkconf[dot]net, see ibid., 34.

in a non-native Spanish speaker'.[43] They give three such slang examples: the repeated use of the word 'careto,' slang for 'face' or 'mask'; the name of an encryption key stored in the configuration files, 'Caguen1aMar', which is probably a contraction of *Me cago en la mar*, identified by Kaspersky's staff as Spanish for 'f—'; and the use of the file path

```
c:\Dev\CaretoPruebas3.0\release32\CDllUninstall32.pdb
```

containing the word *pruebas*, which means *test* in Spanish.

Mistakes are often revealing. Errors can directly reveal information an intruder wanted to keep hidden, such as a name of a person or file, a true IP address, an old email address, or a giveaway comment within the code. Two recent examples are prominent. First, the operator of Silk Road, a site known for facilitating illegal drug sales, used the same username for a web post marketing his illicit enterprise and for a post years earlier seeking technical help. The latter post included an email address with his real name, an obvious clue for investigators.[44] Second, Hector Xavier 'Sabu' Monsegur, one of the leaders of the hacking collective Anonymous, once forgot to log in to the anonymising service Tor before logging into the Anonymous chat system compromised by the FBI, revealing his true IP address.[45] Both individuals have been arrested. Mistakes can be valuable clues even when they do not directly reveal information. For example, frequent typos in commonly-used commands provide a general clue to sophistication. As a rule of thumb, organisations that are more bureaucratic in nature, with more experienced operators and standardised procedures, are less likely to make mistakes than lone activists.[46]

Stealth, ironically, can also be revealing. In any operation, there is a trade-off between speed and stealth that can lay bare clues. Anti-forensic activity — steps designed to evade detection and later investigation — is imperfect and time-consuming. An attacker's use of anti-forensics can reveal intentions, fear of reprisal, and sophistication. Some anti-forensic behaviour is common, fairly easy, and directly linked to mission success. For instance: attackers may encrypt the pilfered data before they exfiltrate it to thwart automated defensive systems that look for valuable files leaving the network. Other anti-forensic behaviours are harder and much less common, for instance using tools to modify timestamps in log files in order to make after the fact

[43]Ibid., 46.

[44]Nate Anderson, and Cyrus Farivar, 'How the feds took down the Dread Pirate Roberts', *Ars Technica*, 3 Oct. 2013.

[45]John Leyden, 'The one tiny slip that put LulzSec chief Sabu in the FBI's pocket', *The Register*, 7 March 2012.

[46]Dan Verton, *Confessions of Teenage Hackers* (New York: McGraw Hill 2000), 83.

investigation more difficult. Caution is hard to measure. But watching attackers' attempts to cover their tracks can be highly insightful.

Some states subject their executive branches, their military, and their intelligence agencies to legal oversight. This means that collection and especially sabotage operations will have to be approved by a legal department, which often restrict the activity. Signs of these restrictions are sometimes visible through technical analysis and can inform attribution. Former counter-terrorism and cyber security official Richard Clarke noted that he thought that Stuxnet 'very much had the feel to it of having been written by or governed by a team of Washington lawyers' because of its target verification procedures designed to minimise collateral damage.[47]

The operational layer of the attribution process is designed to synthesise information from a variety of disparate sources. These include information from the technical layer, non-technical analyses, and information on the geopolitical context. Analysts functioning on the operational layer develop competing hypotheses to explain the incident.

Computer network exploitation requires preparation. Analysing the abilities required to breach a specific network can be a useful clue in the attribution process. The Stuxnet attack on Iran's heavily-guarded nuclear enrichment facility was highly labour-intensive. The malware's payload required superb target-specific information, for instance hard-to-get details about specific frequency-converter drives used to control rotational speeds of motors; about the detailed technical parameters of the Iranian IR-1 centrifuges in Natanz; or about the resonance-inducing critical input frequency for the specific configuration of these machines.[48] Stuxnet also used an unprecedented number of zero days, four or five, and exhibited the first-ever rootkit for a programmable logic controller (used to control industrial machinery).[49] These characteristics drastically limited the number of possible perpetrators. Other preparations include target reconnaissance and payload testing capabilities. Again Stuxnet is a useful example: the attack reprogrammed a complex target system to achieve a kinetic effect. This required advance testing.[50] The testing environment would have to use IR-1 centrifuges. Such machinery can be expensive and hard to obtain.

[47]Ron Rosenbaum, 'Cassandra Syndrome', *Smithsonian Magazine* 43/1, (April 2012), 12.

[48]Ivanka Barzashka, 'Are Cyber-Weapons Effective?', *RUSI Journal*, 158/2 (April/May 2013), 48–56, 51.

[49]Kim Zetter, 'How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History', *Wired Magazine*, 11 July 2011.

[50]William Broad *et al.*, 'Israeli Test on Worm Called Crucial in Iran Nuclear Delay', *New York Times*, 15 Jan. 2011.

No non-state actor, and indeed few governments, would likely have the capability to test Stuxnet, let alone build and deploy it. This further narrows the possibilities.

Offenses vary in scope. They may be isolated incidents against one target or they can be part of a larger campaign that stretches across various victims, possibly for a long period of time and over a larger geographical area. Those who conduct these multi-stage campaign operations are often referred to as Advanced Persistent Threats. These groups often maintain their tactics, infrastructure, and general target set from one operation to the next, so the concept of an Advanced Persistent Threat is a key heuristic in the attribution process. A notable example is a group known as APT1, or the Comment Crew, believed to be comprised of Chinese hackers and known for its sloppy re-use of social engineering tactics and specific infrastructure.[51] Another, probably more experienced group, is tracked by the security firm Symantec in an effort known as The Elderwood Project. This group is known for its unusually frequent use of rare vulnerabilities, its reliance on malicious code known as Hydraq, and its focus on targeting the defence, information technology, and non-profit sectors.[52] How a series of clustered events becomes an Advanced Persistent Threat depends on methodology. The methodologies for making distinctions about scope vary across the information security community. As a result, one company or intelligence agency may conclude one campaign is smaller or larger than another group of analysts might.[53]

Some attacks have multiple stages. Different stages may target different victims, which can hamper reconstructing the campaign design. One breach, in other words, may be merely a stage that enables a larger, more complex breach. The elements of such a large attack can diverge significantly, making it difficult to put the pieces together. For example, the 2011 hack on security firm RSA, itself a multi-stage operation, was part of a larger operation. The breach compromised the SecurID system sold by RSA and widely used by governments and businesses. A follow-on intrusion at Lockheed Martin reportedly leveraged the compromise of SecurID to gain entry.[54] Perhaps an even more elaborate staged attack was the case of DigiNotar. A self-identified pseudonymous Iranian hacker, 'Comodohacker,' first broke into DigiNotar, a Dutch government-affiliated certificate authority, which verifies web servers. Once he had compromised the certificate authority, he

[51] *APT1,* Alexandria, VA: Mandiant, 18 Feb. 2013.

[52] Gavin O'Gorman and Geoff McDonald, *The Elderwood Project,* Symantec, 6 Sept. 2012.

[53] Author conversations with various analysts over the spring and summer of 2014 in Toronto, London, and Washington.

[54] Christopher Drew, 'Stolen Data Is Tracked to Hacking at Lockheed', *New York Times*, 3 June 2011.

issued a significant number of fake certificates, posing as Google and other sites. These certificates then enabled him to intercept the encrypted email traffic of as many as 300,000 unsuspecting Iranians.[55]

Intrusions can evolve. Some campaigns develop over time in ways that do not correspond to pre-planned stages. This development can provide clues to changing political and technical realities and objectives. Stuxnet again provides a noteworthy example. The centrifuge-busting malware came in different variants, noted Ralph Langner, a control system expert who contributed to analysing Stuxnet.[56] These variants used different methods, and they were released at different times. After the main malware was discovered in July 2010, retrospective analysis revealed that the first versions of the trailblazing attack tool were observed as early as November 2005. The earliest version had a different propagation mechanism, did not contain any Microsoft zero-days, and had a working payload against Siemens 417 programmable logic controllers (PLCs) that seemed disabled in later versions.[57] Such shifts in tactics can indicate changing priorities and circumstances.

The geopolitical context of an event can be a tip-off. In hindsight, the geopolitical context of specific incidents may appear obvious: for instance after the DDoS attacks in Estonia in 2007 or during the Georgia War in 2008.[58] But these cases are probably an exception. Interpreting the geopolitical context of an intrusion may require specific regional, historical, and political knowledge about specific actors and their organisation. An example is Gauss, a targeted campaign against Lebanese financial institutions that became public in the summer of 2012.[59] Observers suspected the campaign's rationale was uncovering Hizballah money laundering.[60] Especially for unclaimed and highly targeted breaches, the geopolitical context may limit the number of suspects significantly. Technical analysts are ill-equipped to perform this analysis.

Employees and contractors represent an organisation's greatest strength and greatest risk at the same time. A Verizon review of incidents in 2013 identified insider threats and misuse as one of the most significant risks to organisations. It counted more than 11,000

[55]For a detailed description of the incident, see Thomas Rid, *Cyber War will Not Take Place* (Oxford/New York: OUP 2013), 26–9.
[56]Ralph Langner, 'Stuxnet's Secret Twin', *Foreign Policy*, 19 Nov. 2013.
[57]Geoff McDonald, Liam O' Murchu, Stephen Doherty and Eric Chien, *Stuxnet 0.5: The Missing Link,* Version 1.0, Symantec, 26 Feb. 2013.
[58]Ronald J. Deibert, Rafal Rohozinski and Masashi Crete-Nishihata, 'Cyclones in Cyberspace', *Security Dialogue* 43/1 (2012), 3–24.
[59]*Gauss,* Kaspersky Lab, 9 Aug. 2012.
[60]See David Shamah, 'New virus may be US, Israeli digital strike against Hezbollah', *Times of Israel*, 13 Aug. 2012.

confirmed incidents that involved individuals with privileged access.[61] One of the most costly attacks ever, the Shamoon attack at Saudi Aramco, could have been enabled by an insider.[62] Among the small number of known incidents that involve the successful intentional sabotage of Industrial Control Systems, insiders are the most common cause. Noteworthy cases are the Maroochy Water Breach in Queensland, Australia, in March 2000,[63] and an alleged pipeline incident at Gazprom the same year.[64] A number of control system incidents perpetrated by insiders are likely to have happened since, even if they have never been reported in public. The likelihood that insiders aided a malicious event should be considered higher when the activity required hard-to-get proprietary knowledge, although it should never be excluded from the outset.

On a strategic level, leaders and top analysts are tasked with aggregating the answers to operational questions and drawing meaningful conclusions. The strategic element of the process is at its best when leaders and high-level analysts critically question preliminary analyses, probe for details, and seek alternative explanations.

Cyber attacks are not created equal. The damage caused is one of the most important distinguishing features of a network breach. The damage of a cyber attack, in contrast to offences that involve physical violence, is almost always exceedingly difficult to pin down and to quantify. Damage falls into four broad sets: first, costs can be *direct and immediate*, for instance reduced uptime of servers that causes reduced availability of files, reduced integrity of data, or even hardware that is incapacitated by the intruders. One of the breaches with the highest immediate costs of this kind was the Shamoon attack against Saudi Aramco in August 2012, which incapacitated 30,000 work stations in one go.[65] Second, costs can be *direct and delayed*. Stuxnet manipulated Iranian nuclear centrifuges in such a way that stressed their components. Over a period of months if not years, a code-induced attrition campaign led to deliberate mechanical breakdowns.[66] Costs, third, can also be *indirect and immediate*, for instance reputational

[61]'2014 Data Breach Investigations Report', *Verizon*, 22 April 2014, 23.
[62]Jim Finkle, 'Exclusive: Insiders suspected in Saudi cyber attack,' *Reuters*, 7 Sept. 2012.
[63]Jill Slay and Michael Miller, 'Lessons Learned from the Maroochy Water Breach', in E. Goetz and S. Shenoi (eds.), *Critical Infrastructure Protection*, Vol. 253 (Boston, MA: Springer 2008), 73—82.
[64]Paul Quinn-Judge, 'Cracks in the System,' *Time*, 9 June 2002.
[65]Christopher Bronk and Eneken Tikk, 'The Cyber Attack on Saudi Aramco', *Survival* 55/2 (April–May 2013), 81–96.
[66]Ralph Langner, 'Stuxnet's Secret Twin', *Foreign Policy*, 19 Nov. 2013. For a detailed discussion of Stuxnet, see Kim Zetter, *Countdown to Zero Day* (New York: Crown 2014).

damage or loss of confidentiality. An example is a massive breach at eBay, in which 145 million customer records were compromised.[67] Finally costs can be *indirect and delayed*, for instance a loss of intellectual property that may result in improved market competition once a competitor has been able to utilise the exfiltrated material. One controversial example is the demise of Nortel, a Canadian networking equipment manufacturer.[68] In general, the more indirect and the more delayed the costs, the harder it becomes to quantify them.

The form of the damage may reveal an attacker's intent, especially when properly contextualised on the operational level. Sabotage, as a rule of thumb, tries to maximise direct costs, either openly or clandestinely, whereas collection tries to avoid direct costs for the victim, in order to avoid detection and enable more collection in the future. Of course, the type of target that is damaged also gives clues to intent, as different attackers will prioritise different things.

Intended and actual damage may diverge in two ways. The first possibility is damage was intended but not realised. When Saudi Aramco suffered its major breach in 2012, executives suspected that the attackers had intended but failed to sabotage control systems that run Aramco's oil production. The opposite scenario is that damage was realised but not intended. Computer systems can be complex, and attackers may not know the network's topology. They may thus inadvertently cause damage when performing reconnaissance. Analysts thus must contextualise the damage assessment with other areas of analysis. A cyber attack that causes a minor power outage could be a warning shot, a failed attempt at a major strategic network breach, or an inadvertent result of reconnaissance.

Understanding the rationale of an intrusion is hard but crucial. Knowing an adversary's motivation and behaviour makes mitigating future breaches easier. Such strategic analysis is non-technical by definition. For example, it relies on solid information and analysis from the operational layer on geopolitical context. Against this backdrop, analysis of objectives also requires understanding the priorities of other states, whether they are commercial, military, or economic in nature. All of this can contextualise what a cyber attack was designed to do. It can also provide a clue to an adversary's future action. If an attempted operation failed, understanding why it failed, and what the adversary might do in the future to correct that failure, is helpful for mitigation and response.

Cyber operations are so new that 'firsts' are not uncommon. Analysing these precedents and trying to uncover what they portend

[67]Andrea Peterson, 'eBay asks 145 million users to change passwords after data breach', *Washington Post*, 21 May 2014.

[68]Siobhan Gorman, 'Chinese Hackers Suspected in Long-Term Nortel Breach', *Wall Street Journal*, 14 Feb. 2012.

for the future is not easy: a new method may either be a one-off, or the beginning of a trend. Some may reveal new possibilities, like the programmable logic controller rootkit in Stuxnet that enabled control of industrial control systems. Others may be noteworthy but less significant, like the use of hijacked data centres in the distributed denial service attacks on American banks in the fall of 2012 — a new technical step, but not one of wider strategic importance.[69] Determining if an event sets a meaningful precedent can inform both the attribution process and the response.

Probing the outcome of the attribution process is crucial. The available evidence and preliminary conclusions need testing. Forensic experts are closest to the most tangible evidence, in the form of log files and lines of code. Operational analysts draw on this work alongside other sources. At the strategic level, policy-makers and high-level analysts can provide great benefits to the process as a whole by probing the competing hypotheses produced by the lower levels. Stress-testing the analysis can reveal flimsy assumptions, a lack of imagination, and group-think. Coaxing and probing for additional detail, or for alternative explanations, may require detailed knowledge of the process. This analysis and model is designed to facilitate such probing. If the stakes are high enough, a dedicated red team may even be tasked to go through the entire process again, or to double-check the work of the original team. As Winston Churchill famously said, 'it is always right to probe'.[70]

Part III

Communicating attribution is part of attributing. In complex scenarios, only a small fraction of the attribution process will be visible to senior officials and politicians, and an even smaller fraction to the public. Preparing and managing that portion will determine how an agency's activities are perceived, by the political leadership, by the technical expert community, and by the general public. In many ways, the communication of the process characterises the process for others. Publicising intelligence can harm sources as well as methods. Release decisions are difficult, and officials will often err on the side of caution and secrecy. There are many good reasons for doing so. Yet, perhaps counter-intuitively for those steeped in a culture of secrecy, more openness has three critical benefits: communicating more details means improved credibility, improved attribution, and improved defences.

[69]Nicole Perlroth and Quentin Hardy, 'Bank Hacking Was the Work of Iranians, Officials Say', *New York Times*, 8 Jan. 2013.
[70]Winston S. Churchill, *The Gathering Storm: The Second World War*, Volume 1 (New York: Rosetta Books 2002), 415.

First, *releasing more details will bolster the credibility of both the messenger and the message*. Two recent US examples offer an instructive contrast. On 11 October 2012, the US Department of Defense commented on one of the most high-profile attacks on record. Leon Panetta, then the Pentagon's number one, gave a much-noted speech to business leaders aboard the Intrepid Sea, Air and Space Museum. The venue had a powerful subtext: the museum is on a decommissioned *Essex*-class aircraft carrier, the World War II-tested USS *Intrepid*, floating at a pier in the Hudson River in New York City:

> Over the last two years, the department has made significant investments in forensics to address this problem of attribution, and we are seeing returns on those investments. Potential aggressors should be aware that the United States has the capacity to locate them and hold them accountable for actions that harm America or its interests.[71]

In the speech America's defence secretary mentioned an 'alarming' incident 'that happened two months ago, when a sophisticated virus called "Shamoon" infected computers at the Saudi Arabian state oil company, Aramco.' Panetta gave a few details on the attack's execution, but did not explicitly provide any attributive evidence. Then he mentioned Tehran, a few paragraphs after mentioning the malware: 'Iran has also undertaken a concerted effort to use cyberspace to its advantage.' The international press widely interpreted the speech as a senior US official pointing the finger at Iran. Yet America's most senior defence official at the helm of the world's most sophisticated signals intelligence apparatus merely hinted, and did not reveal any explicit link between Iran and the Aramco attack.

The US government employed a sharply different communication strategy 20 months later. In May 2014, the US Department of Justice (DOJ) took a highly unusual step: it indicted five serving members of a foreign intelligence organisation, PLA Unit 61398, for alleged computer fraud and abuse, damaging a computer, aggravated identity theft, and economic espionage. The document was exceptionally detailed: it outlined, highly unusually, six victim organisations in the Western District of Pennsylvania, the nature and value of the exfiltrated data, as well as the timing of extracting sensitive files. Yet the indictment did not reveal a great amount of attributive evidence. It contained statements such as, 'the co-conspirators used hop points to research victims, send spearphishing emails, store and distribute additional malware, manage

[71]Leon Panetta, 'Remarks on Cybersecurity to the Business Executives for National Security', New York City, Washington DC: Department of Defense, 12 Oct. 2012.

malware, and transfer exfiltrated data'.[72] The subtext was that the government could produce such specific IP addresses, emails, malware samples, and stolen documents, but the indictment itself provided very few forensic details. In this respect the DOJ document was less detailed that Mandiant's APT1 report on the same PLA unit published 15 months earlier. Nevertheless, releasing these details bolstered the government's case and its overall credibility on attribution.

A second reason favours release: *publishing more details will improve attribution itself*. When a case and its details are made public, the quality of attribution is likely to increase. Perhaps the most impressive example is the multi-layered and highly innovative collective analysis of the Stuxnet code: various companies and research institutes analysed the malware and produced a range of highly detailed reports focused on different aspects of the operation.[73] Another example are the more and more detailed reports on Chinese espionage campaigns, partly driven by competition among security companies.[74] As a result, the market for attribution has grown significantly: the most useful and detailed attribution reports that are publicly available are published by companies, not governments. Almost all of the evidence and the examples used in this study come from published company reports. Intelligence agencies have practised attribution for many decades, even centuries. Yet they have done so in relative national isolation, with covert instead of overt competition driving innovation. One consequence of this dynamic is especially noteworthy: the attribution process is not finished with publication, but merely moves into a new stage. This new stage, in turn, may generate new evidence and analysis, and thus require adapting both assessment and outreach campaigns.

The third benefit of openness may be the most significant one. *Making more details public enables better collective defences*. Communication of findings is not just about an individual case, but about improving collective security. For example, a detailed discussion of infrastructure used in an intrusion can enable administrators of other networks to guard specifically against it. Generating new signatures for malicious programs can be similarly beneficial, as they can be downloaded by other administrators and loaded into automated intrusion detection systems. Even absent specific benefits, detailed technical discussion about novel techniques

[72]*United States of America v Wang Dong, Sun Kailiang, Wen Xinyu, Huang Zhenyu, Gu Chunhui,* Criminal Nr 14-118, Erie, PA: US District Court Western District of Pennsylvania, 1 May 2014, 11.

[73]For an overview, see Kim Zetter, *Countdown to Zero Day* (New York: Crown, 2014).

[74]Two of the most notable reports are APT1 and Putter Panda, *APT1,* Alexandria, VA: Mandiant, 18 Feb. 2013, *Putter Panda,* CrowdStrike, 9 June 2014.

used by attackers can better inform investigators in other cases. Providing such indicators and better defences at a price is the business model of cyber security companies. It is an open and important question how governments should react to this dynamic.

Publicity often affects the publicised activity itself. Studying how particular offenders react to the unexpected publicity becomes possible as more attribution reports appear. When Kaspersky Lab, for instance, published its Careto report on 10 February 2014, about a Spanish-language intrusion set, the operation was dismantled 'within one hour'.[75] When the Flame report came out in May 2012, it took the intruders nearly two weeks to shut down the operation.[76] The way an operation is shut down may provide additional attributive clues, for instance whether the shutdown is done professionally, maintaining high levels of operational security, or slowly, possibly indicating that a large bureaucracy had to authorise the decision to shut down the operation. When Duqu, a savvy operation, was revealed, its operators forgot to shred files that had been deleted but recoverable, thus revealing details about the operation.[77] Several of the intrusions that Kaspersky unveiled disappeared after the initial publicity, some faster than others, and some more smoothly than others: Red October disappeared, so did Miniduke and Icefog, all in 2013.[78] The two latter examples are remarkable, because Kaspersky's reports did not identify a suspected offender or even a suspected country; the intruders nevertheless retreated. The handlers of Flame, most notably, started dismantling a highly elaborate command-and-control infrastructure on 14 May 2012, two weeks *before* Kaspersky's report became public, indicating an extraordinary degree of sophistication, possibly even advance warning.[79] When Mandiant published its APT1 report on 18 February 2013, the malicious activity revealed in the highly-publicised report first stopped for 41 days, then remained at lower-than-normal levels until nearly 160 days after exposure.[80] That March the Virginia-based company's websites were nearly overwhelmed by

[75]Costin Raiu, Aleks Gostev, Kurt Baumgartner, Vicente Diaz, Igor Soumenkov, Sergey Mineev, interview with authors, Barcelona, 8 Oct. 2014. See *Unveiling 'Careto'*, Version 1.0, Kaspersky Lab, 6 Feb. 2014.

[76]Alexander Gostev, *The Flame: Questions and Answers,* Kaspersky Lab, 28 May 2012.

[77]Vitaly Kamluk, ' The Mystery of Duqu: Part Six,' Securelist, 30 Nov. 2011.

[78]*'Red October' Diplomatic Cyber Attacks Investigation,* Version 1.0, Kaspersky Lab, 14 Jan. 2013;Costin Raiu, Igor Soumenkov, Kurt Baumgartner and Vitaly Kamluk, *The MiniDuke Mystery,* Kaspersky Lab, 25 Feb. 2013; *The 'Icefog' APT,* Kaspersky Lab, 25 Sept. 2013.

[79]Focus group session with Kaspersky Lab, Barcelona, 8 Oct. email communication with Costin Raiu, 12 Oct. 2014, 11:49 BST.

[80]*Threat Report: Beyond the Breach,* Reston, VA: Mandiant, 18 Feb. 2014, 18.

prolonged denial-of-service attacks emanating from China.[81] Intrusions from China, if often less advanced technically, tend to be unusually persistent, even after an attribution report uncovered sensitive details about an operation.[82]

Public communication finally has to reflect that attribution is gradual, not absolute. Security firms and governments therefore should heed a well-established practice: using words of estimative probability. 'In intelligence, as in other callings, estimating is what you do when you do not know', Sherman Kent, a pioneer of intelligence analysis, wrote in 1968.[83] Estimative language, in Kent's timeless phrase, is 'a mix of fact and judgment'. This mix of fact and judgement is especially relevant in a cyber security context. Estimates are deliberately phrased in a vulnerable way, and therefore open to criticism. The more honest a document is about its limitations of knowledge and about the nature of its estimates, the more credible is its overall analysis. For intelligence estimates are, to quote Kent yet again, 'the next best thing to knowing'.[84]

Conclusion

This study introduced a systematic model for attributing cyber attacks and articulated three core arguments: first, that attribution is an art: no purely technical routine, simple or complex, can formalise, calculate, quantify, or fully automate attribution. High-quality attribution depends on skills, tools, as well as organisational culture: well-run teams, capable individuals, hard-earned experience, and often an initial, hard-to-articulate feeling that 'something is wrong'.[85] The second argument was that attribution is a nuanced and multi-layered process, not a problem that can simply be solved or not be solved. This process requires careful management, training, and leadership. The third argument was that attribution depends on the political stakes. The more severe the consequences of a specific incident, and the higher its damage, the more resources and political capital will a government invest in identifying the perpetrators. Attribution is fundamental: almost any response to a specific offence — law enforcement, diplomatic, or military — requires

[81]Richard Bejtlich, email communication, 11 Oct. 2014, 01:41 BST.
[82]One example is the so-called NetTraveler campaign, which simply moved its command-and-control servers to Hong Kong, then continued operating from there, email communication with Costin Raiu, 12 Oct. 2014, 11:49 BST. See *The NetTraveler,* Kaspersky Lab, 4 June 2013.
[83]Sherman Kent, 'Estimates and Influence', *Studies in Intelligence* 12/3 (Summer 1968), 11–21.
[84]Ibid.
[85]Focus group sessions with analysts from the private and public sectors, Summer 2014.

identifying the offender first. Governments get to decide how to do attribution, and they get to decide when attribution is good enough for action.

Our analysis of the practice of attribution calls into question several commonly held positions in the debate on cyber security. One is that offenders from criminals to spies to saboteurs can cover their traces, stay anonymous online, and hide behind the attribution problem.[86] But attribution is not just possible; it has been happening successfully for a long time. Attackers cannot assume that they can cause serious harm and damage under the veil of anonymity and get away with it. Even if the attribution problem cannot be solved in principle, it can be managed in principle.

A second hackneyed view is that the Internet is taking power away from states and giving it to weak non-state actors, private entities, and criminals; that technology is levelling the playing field.[87] In attribution, the reverse is the case: only states have the resources to open the aperture wide enough to attribute the most sophisticated operations with a high level of certainty. The National Security Agency (NSA) and the Government Communication Headquarters (GCHQ) leaks of 2013 have not only shed light on this dynamic; the revelations have, ironically, strengthened the attributive credibility of these agencies in the eyes of many outsiders predisposed to overestimate their capabilities.

A third common assumption is that the most industrialised and connected countries are the most vulnerable countries, while less advanced and thus less vulnerable countries have an advantage.[88] Attribution again reverses this logic: the larger a government's technical prowess, and the larger the pool of talent and skills at its disposal, the higher will be that state's ability to hide its own covert operations, uncover others, and respond accordingly.

Yet another staple of the debate challenged by this analysis is that the Internet is an 'offence-dominated' environment.[89] Intruders, this view holds, have a structural advantage over defenders, and that advantage is

[86]Perhaps the best articulation of this view is Richard Clayton, *Anonymity and Traceability in Cyberspace*, Vol. 653, *Technical Report* (Cambridge: Univ. of Cambridge Computer Laboratory 2005).

[87]See, for instance, Joseph S. Nye, *Cyber Power* (Fort Belvoir, VA: Defense Technical Information Center 2010).

[88]For instance Michael McConnell, 'Cyberwar is the New Atomic Age', *New Perspectives Quarterly* 26/3 (Summer 2009), 72–7.

[89]For one of the first articulations, see John Arquilla and David Ronfeldt, *The Advent of Netwar* (Santa Monica, CA: RAND 1996), 94; also Department of Defense, *Cyberspace Policy Report*, Nov. 2011, 2.

rooted in the Internet's technical architecture. The defence has to get it right all the time; the offence has to get it right only once. In attribution, again, the opposite is the case: an intruder needs to make only one mistake, and the defender's forensic analysis could find the missing forensic clue to uncover an operation.

Nevertheless, a closer look the limits of attribution is crucial. The first serious limitation concerns resources, especially skill and capabilities. *The quality of attribution is a function of available resources*. Top-end forensic skills as well as organisational experience in complex operations remain scarce, even in a fast-growing international cyber security market. The less resources are available for attribution, the lower will be its quality. The second serious limitation is time: *The quality of attribution is a function of the available time*. Analysing a well-executed operation in a narrow timeframe will be a significant challenge even for the most professional and best resourced teams, firms, and agencies. In serious cases, when high-level decisions will have to be made under pressure, the speed of political developments may outpace the speed of the attribution process. The less time is available for attribution, the lower will be its quality.

A third important limitation concerns the adversary's behaviour: *the quality of attribution is a function of the adversary's sophistication*. The most generally convincing evidence that was published in the examined cases is a result of some operator making a mistake, or not considering the forensic implications of using specific methods. Sophisticated adversaries are likely to have elaborate operational security in place to minimise and obfuscate the forensic traces they leave behind. This makes uncovering evidence from multiple sources, and therefore attribution, harder. The silver lining is that adversaries reliably make mistakes. The perfect cyber attack is as elusive as the perfect crime. Nevertheless: the higher the sophistication of the adversary, the longer attribution will take and the more difficult it will be.

Attribution is likely to retain its core features well into the future. The web has evolved drastically since 1999; but the Internet has not. The net's underlying architecture is changing only slowly. Hence attribution is changing slowly as well — but it is evolving, and it is evolving in a contradictory fashion. On one hand, attribution is getting *easier*. Better intrusion detection systems could identify breaches in real-time, utilising more data faster. More adaptive networks could raise the costs of offensive action, thus removing clutter and freeing up resources to better identify high-profile breaches. More cyber crime could prompt improved law-enforcement cooperation even among unfriendly states, thus making state-on-state espionage both harder to hide and politically more costly.

But attribution is also getting *harder*. Attackers learn from publicised mistakes. The rising use of strong cryptography is creating forensic problems and limiting the utility of bulk-collection. Hype and crises could obstruct nuanced communication. Attribution fatigue may set in. Indeed, absent meaningful consequences, states and non-state actors may simply lose their fear of getting caught, as a lax de-facto norm of negligible consequences emerges. Ironically this could mean that non-democratic states become less concerned about getting caught than publicly accountable liberal democracies. Thus the discussion returns to our central starting point: the attribution process, a techno-political problem, is what states make of it — by investing time, resources, political capital, and by trying to outcompete their adversaries.

The central limitations of attribution also point to the limitations of this study. Some of the most significant attribution work remains hidden and classified in various countries. In the future, government agencies or security companies may develop additional tools — or make tools public — that could open new angles of attribution. Some signals intelligence capabilities may already increase the aperture of the attribution process: even highly sophisticated adversaries who make few or no mistakes could theoretically be uncovered. This study did not benefit from insight into developments and capabilities that are not available in the public domain. Nevertheless, this analysis is likely to have a significant shelf-life. The core variables of attribution have remained remarkably constant since the first major state-on-state campaign was discovered, MOONLIGHT MAZE in 1998.

This text aimed to make progress towards two goals. The first is increasing the quality of bureaucratic output. Time-constraints put significant stress on high-quality attribution, especially when the stakes are high. The model is therefore designed to ensure quality and make attribution more efficient and resilient: the detailed graph, we hope, will help senior leaders in public administration as well as parliament to understand how evidence was generated, to ask better-informed questions, to detect perception bias, and thus to probe and improve the output. At the same time the model will allow analysts at all levels to place their contribution into the larger context of a complex political process. The second goal is broader: increasing the quality of the public discussion. The quality of the wider cyber security debate has been disappointingly low. This includes the technology coverage in some of the world's best news outlets. The scholarly literature in political science and international relations would significantly benefit from more attention to technical details and limitations. We hope that the Q will contribute to raising these standards.

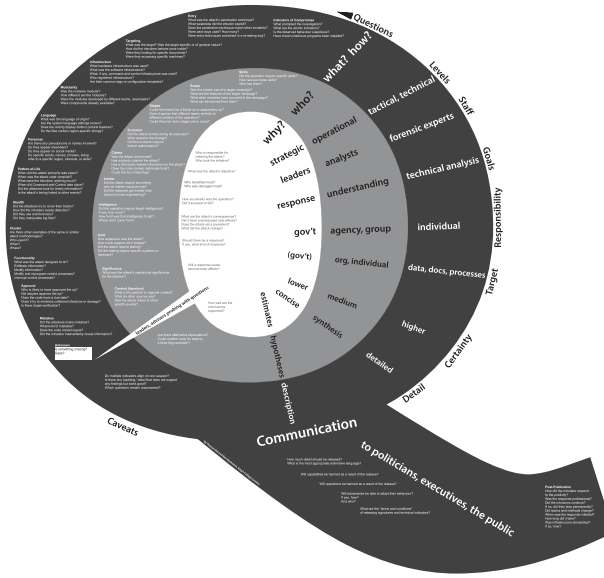

Figure 3. The Q Model, detailed view.

Annex

The Q is designed as a a map of the attribution process: it allows individuals without a technical background to look at the bird's-eye view of attribution in low resolution. It allows scholars as well as politicians or executives to zoom into significant technical detail and enter meaningful conversations with technical experts. Conversely the model enables forensic analysts to appreciate the strategic and political context.

In this article's online PDF version, the image above has unlimited resolution. A separate graph is at http://dx.doi.10.1080/01402390.2014.977382.

The best format is A0, print measures 841mm × 1,189mm.

Acknowledgements

The authors wish to thank Dmitri Alperovitch, Ross Anderson, James Allen, Richard Bejtlich, Kurt Baumgartner, Kristen Dennesen, Brandon Dixon, Vicente Diaz, Alexander Gostev, Mike Goodman, Bob Gourley, Clement Guitton, Nathaniel Hartley, Jason Healey, Eli Jellenc, Robert Lee, Joe Maiolo, Sergei Mineev, Daniel Moore, Ned Moran, David Omand, Costin Raiu, Marcus Sachs, Igor Soumenkov, Jen Weedon, two anonymous reviewers, and members of the intelligence and security community in the United Kingdom and the United States who have to remain unnamed. Several companies provided valuable insights, especially CrowdStrike,

FireEye, Kaspersky Lab, and Booz Allen Hamilton. The views expressed in this paper are solely those of the authors; potential mistakes are their responsibility alone.

Notes on Contributors

Thomas Rid is a professor in the Department of War Studies at King's College London. He is author of *Cyber War Will Not Take Place* (Oxford University Press/Hurst, 2013).

Ben Buchanan is a PhD Candidate in War Studies and a Marshall Scholar. He is also a certified computer forensic analyst.

Bibliography

Anderson, Nate and Cyrus Farivar, 'How the feds took down the Dread Pirate Roberts', *Ars Technica*, 3 Oct. 2013.

APT1, Alexandria, VA: Mandiant, 18 Feb. 2013.

Arquilla, John and David Ronfeldt, *The Advent of Netwar* (Santa Monica, CA: RAND 1996).

Axelrod, Robert and Rumen Iliev, 'Timing of cyber conflict' *PNAS*. 111/4 (28 Jan. 2014), 1298–303.

Barzashka, Ivanka, 'Are Cyber-Weapons Effective?,' *RUSI Journal* 158/2 (April/May 2013), 48–56.

Betts, Richard K, 'Analysis, War, and Decision: Why Intelligence Failures Are Inevitable', *World Politics*. 31/1 (Oct. 1978), 61–89

Betz, David and Tim Stevens, *Cyberspace and the State*, Adelphi Series (London: IISS/Routledge 2011)

Boebert, W. Earl, 'A Survey of Challenges in Attribution', in Committee on Deterring Cyberattacks (ed.), *Proceedings of a Workshop on Deterring Cyberattacks*. (Washington DC: National Academies Press 2011), 41–54.

Brenner, Susan, '"At Light Speed": Attribution and Response to Cybercrime/Terrorism/Warfare', *Journal of Criminal Law & Criminology*, 97/2 (2007), 379–475.

Bronk, Christopher and Eneken Tikk, 'The Cyber Attack on Saudi Aramco', *Survival* 55/2 (April–May 2013), 81–96.

Caltagirone, Sergio, Andrew Pendergast and Christopher Betz, *The Diamond Model of Intrusion Analysis*, ADA586960 (Hanover, MD: Center for Cyber Threat Intelligence and Threat Research 5 July 2013)

Churchill, Winston S., *The Gathering Storm: The Second World War*, Volume 1, (New York: Rosetta Books 2002).

Clark, David D. and Susan Landau, 'Untangling Attribution', in Committee on Deterring Cyberattacks (ed.), in *Proceedings of a Workshop on Deterring Cyberattacks*, (Washington DC: National Academies Press 2011), 25–40.

Clausewitz, Carl von, *On War*, translated by Michael Howard and Peter Paret (Princeton UP 1976).

Clayton, Richard, *Anonymity and Traceability in Cyberspace*, Vol. 653, *Technical Report* (Cambridge: Univ. of Cambridge Computer Laboratory 2005).

Deibert, Ronald J., Rafal Rohozinski and Masashi Crete-Nishihata, 'Cyclones in Cyberspace', *Security Dialogue* 43/1 (2012), 3–24.

Department of Defense, *Cyberspace Policy Report*, Nov. 2011.

Gauss, Kaspersky Lab, 9 Aug. 2012.

Global Threat Report, Arlington, VA: CrowdStrike, 22 Jan. 2014.

Goldsmith, Jack and Tim Wu, *Who Controls the Internet? Illusions of a Borderless World* (Oxford: OUP 2006).

Gostev, Alexander, *The Flame: Questions and Answers*, Kaspersky Lab, 28 May 2012.
Hartley, Nathaniel, *Hat-tribution to PLA Unit 61486* (CrowdStrike, 9 June 2014).
Healey, Jason, *A Fierce Domain* (Washington DC: The Atlantic Council 2013).
Hutchins, Eric M., Michael J. Cloppert and Rohan M. Amin, *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains* (Bethesda, MD: Lockheed Martin Corporation 2010).
Kent, Sherman, 'Estimates and Influence', *Studies in Intelligence* 12, (Summer 1968), 11–21.
Knake, Robert K., 'Untangling Attribution: Moving to Accountability in Cyberspace, Planning for the Future of Cyber Attack', Washington DC: Subcommittee on Technology and Innovation, 111th Congress, 15 July 2010.
Langner, Ralph, 'Stuxnet's Secret Twin', *Foreign Policy*, 19 Nov. 2013.
Libicki, Martin, *Cyberdeterrence and Cyberwar* (Santa Monica, CA: RAND Corporation, 2009).
Lindsay, Jon R., 'Stuxnet and the Limits of Cyber Warfare', *Security Studies* 22/3, (2013), 365–404.
Lupovici, Amir, 'The "Attribution Problem" and the Social Construction of 'Violence"'*International Studies Perspectives* 2014, 1–21.
McConnell, Michael, 'Cyberwar is the New Atomic Age' *New Perspectives Quarterly* 26/3 (Summer 2009), 72–7.
McDonald, Geoff, Liam O Murchu, Stephen Doherty and Eric Chien, *Stuxnet 0.5: The Missing Link*, Version 1.0, Symantec, 26 Feb. 2013.
Moran, Ned and James Bennett, *Supply Chain Analysis: From Quartermaster to Sun-shop*, FireEye Labs, 11 Nov. 2013.
Moran, Ned, Joshua Homan and Mike Scott, *Operation Poisoned Hurricane*, FireEye, 6 Aug. 2014.
Nye, Joseph S., *Cyber Power* (Fort Belvoir, VA: Defense Technical Information Center 2010).
O'Gorman, Gavin and Geoff McDonald, *The Elderwood Project*, Symantec, 6 Sept. 2012.
Panetta, Leon, 'Remarks on Cybersecurity to the Business Executives for National Security', New York City, Washington DC: Department of Defense, 12 October 2012.
Putter Panda, CrowdStrike, 9 June 2014.
Raiu, Costin, Igor Soumenkov, Kurt Baumgartner and Vitaly Kamluk, *The MiniDuke Mystery*, Kaspersky Lab, 25 Feb. 2013.
'Red October' Diplomatic Cyber Attacks Investigation, Version 1.0, Kaspersky Lab, 14 January 2013
Rid, Thomas, *Cyber War will Not Take Place* (Oxford/New York: OUP 2013).
Rivner, Uri, *Anatomy of an Attack*, RSA, 1 April 2011.
Roscini, Marco, *Cyber Operations and the Use of Force in International Law* (Oxford: OUP 2014).
Rosenbaum, Ron, 'Cassandra Syndrome', *Smithsonian Magazine* 43/1, (April 2012), 12.
Schoen, Seth and Eva Galperin, 'Iranian Man-in-the-Middle Attack Against Google Demonstrates Dangerous Weakness of Certificate Authorities', *Electronic Frontier Foundation*, 29 Aug. 2011.
Singer, P.W. and Allan Friedman, *Cybersecurity and Cyberwar* (New York/ Oxford: OUP 2014).
Slay, Jill and Michael Miller, 'Lessons Learned from the Maroochy Water Breach', in E. Goetz and S. Shenoi (ed.), *Critical Infrastructure Protection*, Vol. 253, (Boston, MA: Springer 2008), 73–82.
Stoll, Clifford, *The Cuckoo's Egg* (New York: Doubleday 1989).
The Epic Turla Operation, Kaspersky Lab, 7 Aug. 2014.
The 'Icefog' APT, Kaspersky Lab, 25 Sept. 2013.
The Nettraveler, Kaspersky Lab, 4 June 2013.
Threat Report: Beyond the Breach, Reston, VA: Mandiant, 18 Feb. 2014.
Tsagourias, Nicholas, 'Cyber Attacks, Self-Defence and the Problem of Attribution' *Journal of Conflict & Security Law* 17/2 (2013), 229–44.
United States of America v Wang Dong, Sun Kailiang, Wen Xinyu, Huang Zhenyu, Gu Chunhui, Criminal Nr 14-118, Erie, PA: US District Court Western District of Pennsylvania, 1 May 2014.
Unveiling 'Careto', Version 1.0, Kaspersky Lab, 6 Feb. 2014.
Verton, Dan, *Confessions of Teenage Hackers* (New York: McGraw Hill 2000).

Waxman, Matthew C., 'Cyber-Attacks and the Use of Force', *The Yale Journal of International Law* 36 (2011), 421–59.
Weaver, Nicholas, 'A Close Look at the NSA's Most Powerful Internet Attack Tool', *Wired*, 13 March 2014.
Wheeler, David A. and Gregory N. Larsen, *Techniques for Cyber Attack Attribution* (Alexandria, VA: Institute for Defense Analysis 2003).
W32.Duqu, Version 1.4, Symantec, 23 Nov. 2011.
Zetter, Kim, 'How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History', *Wired Magazine*, 11 July 2011.
Zetter, Kim, *Countdown to Zero Day* (New York: Crown 2014).