International Accountability Mechanisms: Political and Legal Feasibility

13 January 2020, Chatham House, London

Meeting summary prepared by Jack Kenny

The workshop sought to investigate the legal and political backdrop to discussions surrounding the question of developing new attribution or accountability mechanisms for international cyber operations. The existing international rules and institutions do not fully capture the special challenges posed by cyber operations in terms of their modus operandi (e.g. extensive reliance on private actors), the related forensics (e.g. decoys, time-delayed logic bombs, etc.), and the general lack of transparency of the field, which as a result is characterized by low levels of norm-compliance and accountability. The workshop considers some of the central legal and accountability challenges by examining the adequacy of the international rules of state responsibility to address cyber operations, the increased tendency to develop collective attribution statements, sanctions and the reasons for opposition or support by different groups of states for developing new attribution institutions. On the basis of such discussions, questions of feasibility, desirability and possible contours of a new attribution or accountability mechanism for international cyber operations will be revisited.

Presenters: Yael Ronen, Isabella Brunner, Dany Efrony, Jack Kenny

Discussion of presentations by participants at this meeting were held under the Chatham House Rules.

Panel 1: International Responsibility for Cyber Operations

Chair: Yuval Shany

Presentation by Yael Ronen, Evidentiary Dimensions of Attributing Unlawful Cyber Operations to States

The paper on which this presentation was based addresses evidentiary issues in attributing unlawful cyber operations to states. It discusses three main issues: the burden of proof, the standard of proof and evidence or the means of proof. There is very little law in terms of black letter treaties and

jurisprudence is not very clear on the matter, so the paper is best understood as a review of practice which is surprisingly quite uniform, but not to the extent that there is crystallizing law. The paper looks at the International Court of Justice (ICJ) as a model because it is an interstate dispute settlement mechanism, but also considers practice in World Trade Organization (WTO) and the Iran-US claims tribunal, as well as human rights tribunals, with the caveat that these work differently.

There is some uniformity in the approaches of various tribunals, which may be predominantly a result of the relationships between tribunals. Rules of evidence are built for the specific bodies and purpose that they need to serve. Different bodies will have different evidentiary rules depending on considerations including whether there are power disparities between the parties, when the level of technicality required for resolving the dispute is different or where the distance between the breach and time the dispute is resolved is different. As for the burden of proof, the party relying on the fact bears the burden of proving it. While this is the common rule, it is not the only one. For example, a model may impose an obligation on the tribunal to establish the facts in the most accurate way possible, in which case the burden should be on the party that has access to that information or is in the position to provide the best information. The purpose of a tribunal may be a mindset of victory versus loss, or the adversarial model, not speaking of the adversarial versus inquisitorial model (both focus on who is going to win). Even where there is a burden on the parties, there is still an obligation on states to cooperate in the administration of justice, which may require states to provide information that is in their exclusive control. For example, a state might not be obligated in terms of the burden of proof, but they might have a soft law obligation to provide information (Pulp Mills 2010). The burden of proof commonly rests with the party wishing to rely on the fact in question, and within the same process at different points different parties will have different burdens depending on who wishes to invoke which fact.

There have been proposals to reverse the burden of proof. The one most commonly mentioned goes back to the Corfu Channel case, where when a state is in exclusive control over information it might bear the burden even if it is the respondent with regard to that fact. This was rejected both at the Corfu Channel case in 1949 and the later in the Avena case in 2004. It has also been rejected by other tribunals. Another proposal to shift the burden of proof is in the context of negative facts, the idea being it is easier for a state to prove compliance with the law by proving what it has done than for the other state or party to establish negative facts. This was applied by the ICJ in the Diallo case (2010) between Guinea and Congo. Sometimes it has been suggested that the human rights tribunal has moved closer to reversing the burden of proof, but most of these cases involved enforced disappearances, where the burden is on the state to provide a credible explanation of what happened to that person. Rather than a shift of the burden of proof, this should be more accurately regarded as a presumption that a state is in a position to rebut. As this is applied in human rights tribunals in a specific context it does not change the big picture of where we stand with regards to the burden of proof in general.

We are faced with the question of how we transfer these findings to the cyber context and whether we need to make any changes. An international attribution mechanism resolving issues of attribution would operate within the blame allocation realm. There is no particular reason to diverge from the regular rule of burden of proof. The respondent state is often in exclusive control of information, and this appears in the relevant discourse, though in many cases the cyber operation takes place in various states, and so those states may be in control of some or all of the information. Literature on reversing the burden of proof often quickly evolves into proposals to change the rules of attribution. Usually these proposals involve either adding or adapting the effective control and overall control criteria with new terms, such as virtual control. The main problem with reversing the burden of proof is that it would require the respondent state to prove a negative fact, and so it would solve one problem while creating another. It does not seem feasible to have two sets of rules as a result of special rules that would apply specifically to cyber operations.

Regarding the standard of proof, or degree of certainty a tribunal or dispute mechanism should have that a factual allegation is correct, a preliminary issue is whether the standard for proving the facts for an attribution should be the same as the standard for proving facts related to the breach of primary obligation. It is not entirely obvious that they should be the same. Proving the breach is proving the primary rule, while proving the attribution is proving the secondary rule, and those are different sets of rules. There is almost no reference to this in any case law or literature. The little that the ICJ has said about the standard of proof very often relates to attribution rather than proving the breach itself. Often the issue of a breach is not so much in dispute; rather, the question is who is responsible and was there a good reason for carrying out the act. The court is silent about whether it is dealing with the attribution rather than dealing with the breach itself. The one exception is the Bosnian Genocide case, where the court stated that the same standard of proof would apply to attribution as to the breach. The Ethiopia Eritrea Claims Court Commission, in two decisions, implies that the rules should not necessarily be the same, though in practical terms it applied the same rules. The travaux preparatoires of the International Law Commission (ILC) seem to distinguish the two sets of rules, but since they explicitly do not deal with primary rules it is difficult to deduce anything concrete. The assumption that there is a scale of standards of proof originates from the Corfu Channel case. The scale of standards is not unique to international law; it exists in domestic systems, too: in some legal systems the standards of proof are higher for criminal law than civil law.

The ICJ uses a large range of terminology with respect to proof, including “too improbable,” “consistent with the probabilities,” “proof to the courts satisfaction,” “sufficient certainty,” “a degree of certainty,” “decisive legal proof,” “firm conclusion,” “conclusive evidence,” “evidence that is fully conclusive” (this solely regarding cases involving state responsibility, and not proof for other purposes). In the Corfu Channel case, which entailed a British claim that the Albanian government had colluded with the Yugoslavian government in positioning underwater mines, the court said that a charge of such exceptional gravity against a state would require a degree of certainty that had not been reached, implying that if an allegation is grave then the standard of proof will be higher. The

relationship between standard of proof and gravity is later referred to in the Oil Platforms and Croatian Genocide cases. It is unclear what the relevant standards are that could determine gravity.

The gravity of an allegation may relate to the norm that is being invoked or the primary rule that is breached. In international law there is not normally a scale of norms, but a distinction is made between peremptory norms and other norms; presumably the violation of a peremptory norm is graver than the violation of a regular norm. The legal consequences of a violation of one is different from the legal consequences of violation of the other. The harsher the consequences, presumably the higher the standard of proof ought to be. Another approach may be taken with respect to the permissible response to a violation: the harsher the potential response, the higher the standard of proof. When there is a use of force there is sometimes a right to respond through force in self-defense, in which case the standard of proof must be higher. Much of this discussion does not relate to a tribunal allocating blame in retrospect, but to a state needing to decide if it can respond with the use of force without first turning to a judicial body. A third possible criterion for gravity might be harm: The greater the harm that is caused, the higher the standard of proof required. Often these three criteria or considerations go together. We regard certain norms as peremptory because their violation causes such grave harm. In the Corfu Channel case, the primary rule breached appears to relate to innocent passage; use of force was not discussed, but the harm that was caused was equivalent to that of a use of force. In Oil Platforms, Judge Higgins refers to the criminal character of the code of conduct as a relevant factor. In the Bosnian Genocide case ,it seems the criminality of the conduct was an issue that led to a discussion of peremptory norms. Different standards of proof beyond a reasonable doubt appear in the Corfu Channel case and the Bosnian Genocide case, correlated to a violation of peremptory norms. The human rights tribunals – European, American, and the Iran-US claims tribunal – use "beyond a reasonable doubt" in a slightly different sense where the acts in question are such that theoretically carry criminal responsibility when attributed to an individual.

The other standard that is mentioned is the clear and convincing evidence standard. This was mentioned explicitly in the Trail Smelter case in 1941 which concerned environmental protection. It is further mentioned in Congo and Uganda in 2005 and the court applied it in Oil Platforms and Nicaragua, where the court did not require proof beyond a reasonable doubt, but did require something significant below that threshold. These cases all concern the use of force which requires the standard of clear and convincing proof. There is also the standard of preponderance, which does not appear in ICJ state responsibility cases but does appear in the Iran-US claims tribunal. This term which originates from American domestic law. The cases that have come before the ICJ may be such that require a higher standard of proof than preponderance, which is understood as a greater probability that something happened than that it did not happen. Cyber operations often do not reach the threshold of a use of force and are more concerned with violations of sovereignty and non-intervention. If these are considered lesser norms than a use of force, then the standard of proof might be slightly lower and so preponderance may be a more relevant standard. There have been proposals to relax the standard of proof when it comes to cyber operations because it may be unfeasible to demand a high standard. Rules of evidence are already a compromise – ideally, we would use a standard of absolute certainty, but in

reality a balance must be found with a certain perception of how things work, what type of things need to be proved, and what means we have to prove them. If it turns out that in cyber operations the issues or means are very different, maybe there is reason to move or adjust this balance. The assumption that we can devise new rules for cyber operations, for example applying a clear and convincing standard for normal kinetic operations and a preponderance standard for cyber operations, as different sets of rules, will inevitably lead to problems distinguishing between cyber and normal operations. The ICJ has rules determining which evidence it will consider more credible, such as something direct, nonpartisan, and that has been tested before. In principle everything is admissible as evidence before the court. Cyber operations might bring to the table more reliance on privileged information and information that has been obtained illegally. Only time will tell whether new rules will be devised in response to this.

Discussion

The discussion began with a clarification of the research project as not necessarily considering an ICJ adversarial process, but rather a mechanism more along the lines of a fact-finding exercise, where an international body assesses the evidence and comes out with a finding or report, akin to the structure of the Organization for the Prohibition of Chemical Weapons (OPCW). Questions were raised by several participants as to whether this changes the understanding of the applicable burden of proof, standards of proof or admissibility of evidence, and how the rules of international responsibility deal with evidence that cannot be universally disclosed, such as intelligence or information that raises national security concerns. The mechanism was also discussed in relation to whether it should perform attribution to non-state actors rather than states. A participant noted a dissonance between the legal thinking of state responsibility and the practice of attribution by private companies that often results in a degree of probability that needs to be translated into standards such as clear and convincing, preponderance etc.

It was agreed that there needs to be more clarity about the term "attribution," whether that is the act of finding out who is responsible or the act of deciding what to do after that is known. If states start taking countermeasures, they must prove that the previous act was unlawful, as they have to prove attribution. As we have not reached that point, we remain at the political level of naming and shaming; but if we move to the legal level, then attribution is important because that decision must be taken before taking countermeasures or sanctions.

Several participants made comments about the distinction between technical, political, and legal tracks of attribution. Political attribution would at least take into account technical attribution, but also pays attention to legal facts. Political attribution can involve work with probabilities, where byit is up to states when they find things clear enough to make an attribution statement. Once you have the legal elements in place then there is a policy decision whether to attribute publicly. To date we have not seen states make legal cases for attribution for cyber operations, instead only political attributions. Several participants questioned whether there was an attribution problem for cyber operations, and what issues such a mechanism sought to solve or improve. If carefully collated, open source material provides a great deal of information that can make significant progress towards performing attribution. Those

involved in responding to the attacks publish details of the attacks to prevent further activity, and due to the complexity sophistication and long running nature of the operations the actors make mistakes which allow them to be identified. Participants questioned whether a UN body would be a productive mechanism.

Panel 2: Collective Attribution for Cyber Operation

Chair: Harriet Moynihan

Presentation by Isabella Brunner: Collective State Attribution for Cyber Operations: An Analysis of Existing Approaches

There is a trend towards collective attribution by states at the international level. Recently there has been increased interest in trying to establish an independent fact-finding body, in academia but also in the private sector. However, states seem to be divided: the Open Ended Working Group (OEWG) appears to include less cyber-capable states that are interested in such a mechanism, but cyber-capable states may not share this interest. The European Union (EU) established the cyber diplomacy toolbox in June 2017 – a set of measures with which the EU can respond to malicious cyber activities including demarches, public statements condemning the actions of a specific state, confidence building measures, political dialogues, or restrictive measures/sanctions. For some measures, attribution would be required where a target state needs to be identified. The EU distinguishes between firm and less firm attribution but there is no further information on what that means, and whether that should be regarded as some sort of evidentiary standard.

The diplomacy toolbox has been applied on three occasions, including the OPCW hack in 2018. For establishing attribution, the EU Intelligence and Situation Centre (INTCEN) plays a key role in cooperation with other EU cyber security agencies where they are tasked to undertake some sort of analysis of information they have gathered themselves or have received by other member states. In certain early papers, uncertainty yardsticks are used to divide probabilities into percentages to make an assessment. It is not clear if the member states will receive information from INTCEN or what the exact basis for sharing information is. INTCEN will not take a decision on attribution but only provide analysis after which it will be up to member states whether to make public attribution statements. One of these measures is the cyber sanctions regime adopted in May 2019, the second regime in the world establishing sanctions against cyber behavior after the US. It imposes entry restrictions and the freezing of funds on certain entities, persons or bodies who are responsible for a cyber-attack under the threshold of an armed attack under Article 51 of the UN Charter. Every sanctions regime contains a list of persons or entities who are subject to restrictions, however, to date, the list has not been used. The listing procedure is a unanimous decision by member states. The listings can be challenged, and the Council is obligated to review the sanctions regime every 12 months. It is clearly stated that targeted restrictive measures should be differentiated from the attribution of responsibility of cyber attacks to a third state.

The attribution or listing is not the same as attribution in international law. The evidentiary standard required to list a person or entity is that the evidence must be sufficient to withstand the examination of the European Court of Justice (ECJ). The court in the past has ordered the delisting of certain individuals where the reasons were not sufficient and stated that excessively vague reasons would be in violation of the so-called obligation to state reasons. When states are undertaking public attribution of cyber operations they tend to be vague. NATO considers cyber-attacks capable of reaching the threshold of armed attack under Article 51 of the UN Charter which would trigger Article 5 of the North Atlantic Treaty. Scholars discuss the possibility of whether Article 4 of the North Atlantic Treaty, which concerns territorial integrity or political independence, may lead to a collective statement condemning malicious cyber behavior.

Capacity building measures and cooperation with the EU on information and intelligence sharing could both assist collective attribution statements by states. The US National Cyber Strategy discussed the cyber deterrence initiative that entails attribution and collective attribution with other states. This US promotion of a collective approach to attribution preceded NotPetya and Wannacry. While the EU system is institutionalized, the US favors ad hoc alliances of collective state attribution. The Guide to Cyber Attribution document published by the US Office of the Director of National Intelligence outlines definitions of high, medium, and low confidence levels of attribution. The Australian position at the OEWG committed to cooperation with international partners to perform joint attribution and discussed the capability to attribute attacks to several levels of categories to actors and states. In conclusion, states are increasingly more active in collective attribution.

OEWG discussions suggest more cyber capable states are not in favor of a UN body but in broadening efforts for performing collective state attribution seeking increased political pressure and more credibility. This attribution should be understood in a wider spectrum of confidence building and capacity measures, and attribution should not necessarily be equated with attribution as understood in the ILC's Articles on State Responsibility.

Discussion

The discussion began with the recognition that in order to stabilize the system of collective attribution, you need a process in which you can have some confidence. Informal intelligence sharing has its limits. There was discussion among the participants over the extent to which a more robust process may compensate for difficulties in disclosing intelligence information. The current ad hoc approach to attribution by groups of states requires a certain relationship of trust that enables states to share and rely on information, and this limits the scope of the group of states that may engage in collective attribution.

Differences were highlighted between political attribution and legal attribution. Where responding with countermeasures, intelligence sharing may not be a sufficiently robust basis. An attribution

mechanism could be useful in order to balance confidentiality and generate trust and some degree of accountability. From the EU's point of view – and Kadi type case law – there is a problem with relying on certain intelligence information and generating legal repercussions from that information. We have not yet observed situations where states are publicly making the case to invoke countermeasures in response to cyber operations. If cyber operations rise to thresholds where primary rules are frequently violated, there may be a need for a more robust evidentiary basis to invoke legal consequences. A participant explained that the standard of evidence for EU cyber sanctions is sufficiently solid evidence and that there is already an established process through other sanctions methods where the Council Legal Service is required to defend them within the ECJ on the basis of evidence submitted to all EU member states. The cyber toolbox has been translated from an existing system that has already faced many of these questions.

The discussion returned to the need to identify the problem that an attribution mechanism seeks to address. There was encouragement that states should adopt a deterrence toolkit, which comprises of a process for deciding whether to publicly attribute cyber operations, and make political signals that they are willing to respond for deterrence purposes. participants recognized the importance of the involvement of private technology firms in addition to attribution statements made by states. Some concerns were raised by about a possible risk of abuse in an attribution mechanism, and it was recognized that any mechanism must be structured in a way that minimizes these abuses. The mechanism would not aim to be suitable for immediate attribution and response to cyber operations. It may serve as a mechanism generating a coordination point not just for collective attribution but also to clarify different evidentiary standards and interpretations of the law. Greater uniformity in these areas may facilitate the collective attributions made by states.

Panel 3: Lessons Learned from Other Mechanisms

Chair: Paul Ducheine

Presentation by Dany Efrony: An International Attribution Mechanism: The UN Group of Governmental Experts (UN GGE)

Advancing responsible state behavior in cyberspace in the context of international security became more complex last year after the UN processes split into two directions when a Russian proposal established the OEWG. A different approach exists under the Budapest Convention countering cybercrime with the open-ended intergovernmental expert group, launched by the European Commission (EC) and supported by the US and other states, whose purpose is to produce a comprehensive international convention countering the use of information and communications technologies for criminal purposes. Looking at existing mechanisms, there is an accountability gap where the international community has not succeeded in establishing a convention to build on to support attribution and punitive measures. In considering some examples of existing mechanisms, we observe the mechanisms that are used in the International Atomic Energy Agency (IAEA), a fact-finding verification mechanism that does not impose attribution, and relies on reports from state parties

according to their policies. They carry out on-site inspections and their reports are submitted to the Security Council (UNSC).

The Organization for the Prohibition of Chemical Weapons (OPCW) has a similar structure, also dealing in fact finding and verifications. The commonalities of these two and a third body, the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO), is that they always rely on consensual decisions. Cooperation is therefore the premise of these mechanisms, and there do not seem to be any cases of the inspections being challenged. The CTBTO is also a fact-finding mechanism, but operates as an attribution mechanism because it is scientific, based on the work of almost two decades of an international group of experts. However, it has not entered into force and will not do so as not enough states have ratified the treaty. A fourth mechanism is the Proliferation Security Initiative (PSI), which is not a mechanism for attribution but a statement by states that are ready to play an active role in enforcing the ban of transferring weapons of mass destruction (WMDs) and materials associated with WMDs. It appears successful because many states have joined, but as noted does not function as an attribution mechanism.

Regarding the proposals for a mechanism for cyber attribution, there is the Multilateral Cyber Attribution and Adjudication Council (MCAAC), initiated by the Atlantic Council, the International Cyberattack Attribution Organization (ICAO), proposed by a Microsoft team of researchers based on the IAEA model, and the Global Cyber Attribution Consortium (GCAC) initiated by RAND. The GCAC is significant since it is the only proposal that rejects the involvement of any states in its work. The MCAAC and the ICAO support the involvement of states and would like to have the involvement of the five permanent members of the UNSC participating in the mechanisms, along with representatives of those states together with private sector, civil societies and academia. The weakness of the MCAAC is that it is impractical, since it requires consensus and an international convention, in which case we would not need these proposed mechanisms in the first place. The MCAAC, ICAO, and GCAC all rely on the private sector, where attribution is mainly technological, but they can use information from academia or civil society that can contribute to supporting this technical attribution.

There are several other proposed mechanisms in early stages of conception. The international cyber court or arbitrage proposed in a blog by Russian researchers for foreign and defense policy also requires consensus and amending the Budapest convention. Regarding standards of proof, they would need a court to decide and every victim state would bring its information to the court. The Global Cyber Attribution Peer-Review Network of the ICT for Peace and the CyberPeace Institute are both located in Switzerland. The ICT for Peace is supported by the Swiss government, which favored the idea of peer review of the contributions of others in the same network to establish attribution.

Consensual regulation by convention of cyberspace is not feasible for reasons already discussed at this conference. The experience from the WMD parallel and from cyberspace indicate that superpowers are

reluctant to relinquish advantages they have in the context of this sensitive and strategic realm. This is the reason they do not accept situations where the mechanism would be an attribution mechanism rather than mere verification. The fact that China and the US have not ratified the CTBTO indicates they want to maintain a strategic advantage in this situation. Attribution is the lynchpin of accountability even in the absence of binding law. It may have an important role to play in filling the gap by using legitimacy as leverage. If we establish such a mechanism its product would be used efficiently as a tool in the form of legitimacy, not only against Russia and China but also against Western states. The focus of this mechanism should be on politically-motivated cyberattacks and not on private sector hacks. If states reach common ground, we can much more easily deal with other threats in cyberspace. In general, technical attribution is not sufficient to attribute responsibility to a state, but at times it could be decisive, and the CTBTO is a good example of this. New and clear definitions are required referring to the applicable standards of proof and in limiting the definition of what operations may be considered espionage. The contribution of the private sector is important but could not and should not replace the states.

Discussion

We do not have a convention on cyber operations, only the nonbinding norms of 2015 UN GGE that has been adopted by the UNGA. A participant discussed the perceived bias by some states in the advancement of norms applicable to cyber operations, and the perceived bias an attribution mechanism might face should such a proposal take shape. Several participants recognized that reaching agreement on key definitions for performing attribution may be difficult when establishing a state-centric mechanism. There was some discussion by participants about the distinction between facts and law, and the manner in which fact-finding mechanisms sometimes make legal determinations. Attribution is a legal condition for state responsibility based on facts.

Presentation by Jack Kenny, Case Studies in the Attribution of Cyber Operations to a State

The paper this presentation is based on provides an overview and analysis of case studies where states have made attributions of cyber operations. The focus of the case studies is on attribution and the methods or modality in which attribution is made. The case studies are identified and compiled from cyber-attacks which are of a significant gravity in relation to debates surrounding the application of primary rules of international law as to be worthy of attention and discussion. The paper does not focus on the lawfulness of these cyber-attacks or the application of primary rules to these attacks and recognizes the limitation that the possible bases for attribution discussed are based on attribution statements made by states and the private sector.

Whether the evidence relied on by those states and private sector actors in making those attribution statements is sufficient to meet the thresholds of attribution under the corresponding International Law

Commission's Articles on State Responsibility often cannot be established from open source information. In thinking about what makes cyber operations different from traditional kinetic operations, there are a number of points to consider. The ICJ in Nicaragua dealt with difficult issues of attribution of the actions of a non-state actor to a state. With cyber operations there are numerous Advanced Persistent Threat (APT) groups, which are referred to with inconsistent nomenclature. These groups have the ability to carry out frequent attacks, sometimes on a continuous basis and often targeting multiple states simultaneously. Attacks are often routed through one or many state territories instantaneously. Geographical evidence is easily manipulated or hidden. Malicious actors can implicate others as being responsible for an attack through "false flag" operations.

The paper identifies six case studies in the attribution of cyber operations by individual states. In each case study, the paper examines the attribution of the cyber operation by a state to a responsible actor, the detail and confidence level of that attribution, private sector support for that attribution and possible bases for attribution in relation to the ILC's Articles on State Responsibility. The relevant bases for attribution under the ILC's Articles on State Responsibility are acts of a State's organs (Article 4), persons or entities exercising elements of governmental authority (Article 5) and conduct that is directed or controlled by the State (Article 8). The paper identifies seven case studies of co-ordinated state attribution of cyber operations, beginning December 2017 with the NotPetya. It examines the various attribution statements made by states and the level of confidence of those statements, reports from the private sector that support those coordinated state attributions and possible bases for attribution in relation to the ILC's Articles on State Responsibility. Generally, statements attributing cyber operations by states have political and legal considerations. States provide little technical analysis or support for their attributions (outside of private sector attribution reports). An exception here may be the US indictments of foreign actors in-absentia, which frequently provide significant detail on the actors behind the malicious cyber operations in question and their links to a state. Where states refer to a violation of international law, which is rare, they are non-specific about which rules of international law may have been violated.

In conducting this research, several state attribution modalities have been noteworthy. The paper identifies four areas from state positions and policy concerning the attribution of cyber operations: The identification of key indicators in how states perform attribution, the use and definition of confidence levels in performing attribution, the classification or categorization of the severity of cyber-attacks (and corresponding response options) and the provision of further information and context for public attribution statements. The private sector plays an important role in the attribution of cyber operations, from reports and investigations in the media to attribution reports and analysis by cybersecurity and technology firms. Media reports often provide information from cited government and industry sources that isn't available elsewhere, and various different private sector attribution reports from a technical or a political approach can offer different insights into aspects of how attribution may be made to a state or non-state actor. Cybersecurity firms also maintain profiles on known APT groups, their suspected attribution to a state and history of targets and associated malware, which can be an invaluable tool of open source information.

From 2017 onwards, there has been a trend among Western states to coordinate attributions of cyber-attacks in an ad hoc manner in an attempt to increase legitimacy and strengthen accountability. Private sector attribution reports can provide indirect support for attribution statements by states. In examining case studies, it is clear that there is a difficulty separating the attribution of a cyber-attack to an APT group and the attribution of the actions of an APT group to a state. States use unclear language in their attribution statements, and often conflate these two steps. There is a difficulty in collating this information from open source information, in so far as it is available. Some of these issues are confounded by the inconsistent naming and identification of APT groups. It is useful when states break down attributions of cyber campaigns into attributions of specific attacks. The UK's attribution of the campaign of indiscriminate and reckless cyber-attacks to the Russian GRU in 2018 is a good example of how this may be done effectively. Divergent views among states over the application of international law norms to cyber operations at the fifth session of the UN GGE and the parallel proceedings of the newly established OEWG highlight the difficulties of states reaching agreement in this area.

There is certainly some skepticism about the interest of states in establishing an independent fact-finding mechanism. The CyberPeace Institute proposed by Microsoft that involves experts from academia, industry and civil society seems a positive step by the private sector towards assistance, accountability and the advancement of rules of international law governing cyber operations. It is difficult to measure the effectiveness of coordinated state attributions as a deterrent, and without involvement from key states a state-centric attribution mechanism along the lines of the OPCW may risk becoming the face of political attributions made by Western states. In light of this and the aforementioned skepticism among states, and after examining the case studies in the paper, it may be that the private sector is most suited to form such a mechanism, and that Microsoft is best placed to do this, though the CyberPeace Institute is still in its early phases and its impact is yet to be determined.

It is clear from these case studies that the private sector provides large amounts of information, that when carefully collated together can offer a detailed picture to assist in attributing cyber-attacks. The private sector possesses significantly more capabilities and resources than states to investigate and respond to cyber operations. A private sector focused mechanism will not preclude states continuing to make ad hoc coordinated attributions but could provide significant support for those attributions. Any attribution mechanism, state-centric or more aligned with the private sector, would do well to adopt and seek to find common ground to develop standardized usage of the modalities or methods identified in the paper, with a focus on transparency and open source reporting. These include key indicators for the performance of attribution, definition and usage of confidence levels in attributions, classification, and categorizations of cyber-attacks, provision of information and context for attributions, and maintaining APT profiles and attack databases. Such a mechanism could also clarify different evidentiary standards and advance the interpretation and understanding of norms of international law to cyber operations.

Discussion

Discussion by participants focused on what a private based model of attribution mechanism might offer over a state-centric mechanism. Some participants felt that such a mechanism could not replace the role of states and the contribution a state-centric mechanism would make, while others felt that a private sector mechanism was more realistic and might help support ad hoc coordinated attributions by states. Participants discussed the importance of other areas of focus including cooperation and capacity building amongst states to help enable states to better respond to cyber-attacks.

Concluding Panel and Discussion: The Political Feasibility of New Mechanisms

Chairs: Paul Ducheine, Yuval Shany

The concluding panel revisited the question of whether there is an attribution problem regarding cyber operations or whether this is a non-issue. Discussions continued over how the technical, political, and legal attribution of cyber operations present different challenges. Legal attribution to a state that could serve as the basis for subsequent legal action such as countermeasures may not be satisfied by this current ad hoc approach to attribution. However, most cyber-attacks to date are low-level attacks below the use of force. The participants continued to consider whether establishing an international attribution mechanism would make sense politically as opposed to the current ad hoc collective attribution trend. States would have to support the mechanism and be prepared to engage in sharing intelligence to make its operation a successful endeavor.

Some participants expressed the view that smaller states with limited technological capacity and political clout would be more interested in supporting such a mechanism than more advanced and powerful states. If a mechanism were established in spite of these concerns, it might play a useful role in coordinating the various approaches to attribution by states, facilitate collective responses, and develop the interpretation of the law, the rules of attribution, and evidentiary standards in relation to cyber operations. However, in light of the current uncertainty over the application of primary rules of international law to cyber operations, this may not be the right point in time to be pursing acceptance of such a mechanism where cyber capable states are seeking to maintain an advantage in operational flexibility.