British Institute of International and Comparative Law

State Responsibility for Cyber Operations: International Law Issues

Event Report

Date: 9 October 2014

Venue: British Institute of International and Comparative Law
Charles Clore House, 17 Russell Square, London WC1B 5JP

Speakers:

- Dr Russell Buchan (University of Sheffield)
- Dr Marco Roscini (Westminster Law School)
- Professor Nicholas Tsagourias (University of Sheffield)

Chair:

- Cathy Adams (Legal Director, Foreign and Commonwealth Office)

With the number and sophistication of cyber-attacks against states showing a significant increase in recent times, the Institute invited experts on the topic to discuss the international law problems related to state responsibility for cyber operations. Speakers covered questions such as the evidentiary and attribution rules applicable to cyber operations, as well as the possible sanctions available against different types of internet-based attacks.

The seminar was chaired by **Cathy Adams** (Legal Director, Foreign and Commonwealth Office) and the panel of speakers included **Professor Nicholas Tsagourias** (University of Sheffield), **Dr Marco Roscini** (Westminster Law School) and **Dr Russell Buchan** (University of Sheffield).

Professor Nicholas Tsagourias initiated the seminar by discussing international responsibility in cyberspace and the problem of attribution. According to international law, a State incurs responsibility if an act or omission is attributed to that state and constitutes a breach of one of its international law obligations. Attribution is about the assignment of an act to a State. Because states are abstract entities that act through physical persons, attribution establishes a link between the act, its physical author and the State.

Attribution is however a very demanding and complicated exercise in general and this is even more so in cyberspace because of the nature of the cyber domain. Three particular characteristics of that domain make attribution extremely difficult. The first is 'anonymity' in that the authors of cyber operations can hide their identity; the second is the possibility of multi-stage action in that computers operated by different persons and placed in different jurisdictions can be used and the third is the speed with which operations can take place. For example, the denial of service attacks on Estonia in 2007 involved a large botnet of approximately 85,000 hijacked computers from around 178 countries.

What this example demonstrates for purposes of attribution is the need to trace back a cyber-act to its source, for example to a computer, to identify the person that operated the computer and even more importantly to identify the real 'mastermind' behind that person and then establish his/her links with a state. It thus transpires that attribution has technical, political, as well as legal aspects, with each aspect feeding into the other.

Regarding the technical aspect of attribution, they concern the forensic identification of the source of a cyber-act. Although technical attribution can yield good results by tracing back the computer or its geolocation it can never be absolutely exact and, moreover, it cannot identify the person that operated the computer or his/her affiliations. For this reason, in addition to forensic investigation, intelligence and information analysis is needed in order to profile the author of the act, provide information about her capabilities and intentions or her links with states or other entities. This makes attribution political. Of course critical to any such assessment is the availability of evidence and its probity.

In discussing the legal aspects of attribution, Professor Tsagourias first presented the attribution standards included in the law of state responsibility and assessed their relevance when applied to the cyber domain. As they, in his opinion, fail to capture the intricacies of the cyber-space, he then proposed a model of responsibility based on an obligation of due diligence and on causation.

1) International law criteria on attribution

There are three main attribution tests in the law of state responsibility: an institutional test, a functional test, and a control based test. According to the institutional test, acts of state organs are attributed to the referent state. Attribution in this case is premised on the juridical status of the physical author of the act. The institutional test includes *de jure* organs, for example the military, but also *de facto* organs. A *de facto* organ is a person or entity that is assimilated to or absorbed in the State apparatus. In the *Nicaragua* case, the International Court of Justice (ICJ) described a *de facto* organ as one that is in a relationship of complete dependence on a

State and of control by that State. It is not always clear what degree of control is needed with the Court in the *Nicaragua* case mentioning 'effective control' as well as 'general control' whereas in the *Bosnia Genocide* case the ICJ spoke of 'strict control' or of 'great degree of control'. At any rate, a high degree of control is required because, as the ICJ said in the *Nicaragua* case 'to equate persons or entities with State organs when they do not have that status under internal law must be exceptional'.

Secondly, according to the functional test, an act is attributed to a State if it is committed by an entity that is empowered by that State to exercise governmental authority or if it is committed by an organ of another State that has been placed at the disposal of the first State.

Thirdly, according to the control test, an act is attributed to a State if it is committed by an individual or a group that have been instructed or directed or acted under the control of a State.

The standard of instructions would attribute to a State a cyber-act committed by persons prompted by a State organ. Instructions establish an ad hoc relationship between the author of the act and the State. For this reason, instructions should be proven in relation to the specific act.

As far as direction is concerned, persons or groups should have been specifically charged by the authorities of a State to commit a particular act, or to carry out a particular task on behalf of the State. With regard to control, international jurisprudence or doctrine seems to wrestle with the question of the requisite degree of control. The ICJ in the *Nicaragua* case spoke of effective control over the wrongful act. Effective control exists when the State either directly influences the commission of the act or enforces its perpetration.

That said, a lower threshold was also introduced but in a different context. The International Criminal Tribunal for the Former Yugoslavia (ICTY) in the *Tadić* case distinguished the case of individuals and unorganised groups where 'effective control' is the requisite standard from that of organised groups where 'overall control' is needed. A State wields overall control over the group 'not only by equipping and financing the group, but also by coordinating or helping in the general planning of its military activity.' However, it is not necessary for the State to issue to the head of the group or to members of the group, instructions for the commission of specific acts contrary to international law.

Overall control alludes to the ability of the State to exert general and direct influence on the group and on its activities which differs from the effective control test which requires control over the specific act. However, in the *Bosnia Genocide* case the ICJ criticised the 'overall control' test and reaffirmed the effective control test in the law of state responsibility.

If we were to apply these standards to the 2007 attacks on Estonia and the 2008 Distributed Denial-of-Service (DDoS) attacks on Georgia, we realise that these acts evade the existing attribution standards. The 2007 attacks on Estonia targeted governmental and banking sites. They took place in the context of a dispute with Russia. Similarly, the attacks against Georgia in 2008 targeted governmental sites and took place in the context of a military confrontation between Georgia and Russia. Both countries politically attributed the attacks to Russia and

identified the Russian Business Network (RBN) and Nashi, a Russian nationalistic organisation, as the coordinator of the attacks. Yet, neither the RBN nor Nashi are *de jure* or *de facto* State organs. They are not designated as State organs by Russian law and they are not in a state of complete dependence to the government. Furthermore, they are not delegated to exercise governmental powers. Neither can it be said that they acted on the instructions of Russia or under its effective control in the sense that Russia directly instructed their specific acts or enforced their commission. As to whether Russia exercised overall control over them, this depends on whether the Russian Business Network and Nashi are organised groups in the sense used by the ICTY in the *Tadić* case. The *Tadić* case concerned military-like groups exhibiting a hierarchical structure and a chain of command. Moreover, overall control requires financing, equipping, coordinating and help in the planning of the operations. The RBN and Nashi are not militarily organised and Russia was not involved in the planning of their operations although some form of coordination may have existed. Moreover, Russia denied any involvement and no further action was taken at a State level implicating Russia. It becomes evident that these attacks cannot be attributed to any State according to the existing attribution criteria with the possible exception perhaps of the overall control criterion. It is reasonable then to ask whether the attribution criteria as they have been formulated and interpreted in the law of state responsibility are under-inclusive or even more crucially, whether they are relevant in cyberspace. In order to answer this question, it is important to explain the context within which these standards developed.

2) Context

Since the law of State responsibility concerns States, the attribution standards determine which acts are State or public acts for which the States can be held responsible. The State is not held responsible for private or non-State conduct. Thus, acts of private actors are attributed to a State only if those private actors are organs or agents of the State or when their actions are subordinated to the State. In doing so, the law serves the interests of States by limiting the scope of State responsibility. Yet, in cyberspace the interpenetration between public and private actors is more acute. Moreover, private actors are often more dominant than States or pursue public agendas.

Secondly, the attribution criteria are historically contingent. They reflect States' experiences of proxy wars fought by non-State actors (NSAs) with conventional weapons. To fight such wars, these non-State actors needed to have access to heavy weaponry that only States could provide by either equipping or financing them. In cyber, however, non-State actors are not entirely dependent on States and they can be self-sufficient. Cyber weapons in the form of viruses, Trojan horses and their equivalents can be invented by non-State actors, they are easy to acquire and they are quite inexpensive.

Thirdly, the requirement of organisation for groups reflects a traditional view of military-like organisation according to which there needs to be some form of hierarchy, internal regulations and the capacity to launch coordinated operations. Although a cyber-group can act in a coordinated manner or take orders from a virtual leadership, hierarchies and the chain of command may be very diffused. Cyber groups can be social networks often with infinite membership. Another element which may be missing in cyber groups but is important

in traditional definitions of groups is that of the power to enforce the law. This may be difficult in a cyber-group where physical control over members is lacking.

Finally, the fact that cyber actors can maintain their anonymity or 'spoof' attacks or divert attacks through different routes poses a challenge to the current attribution criteria that are based on conventional understandings of physicality where embodied entities operate through physical acts in the real world.

In light of the above, it is not unreasonable to say that the attribution standards should adapt to the exigencies of cyberspace. Some degree of flexibility in the attribution standards has been recognised in the law of international responsibility. For example, according to Article 8 of the 2001 Articles on the Responsibilities of States for International Wrongful Acts (ASR), the standard of control should apply with a degree of flexibility. Article 55 ASR also acknowledges the existence of special regimes which may have their own attribution criteria. The ICJ, on its part, did not deny the existence of different attribution standards for different situations. As it put it in the *Bosnia Genocide* case, 'logic does not require the same test to be adopted in resolving issues which are very different in nature.' That change happened in the use of force regime. When states realised that the attribution criteria left them with no effective remedy against devastating terrorist attacks launched by NSAs from the territory of a State which tolerated them or was unwilling to prevent such attacks, they introduced the attribution test of toleration or unwillingness. According to this standard, if a State fails in its duty of due diligence to prevent or suppress terrorist activities on its territory and that failure causes an armed attack to be committed by a NSA against another State, the former State is responsible for the attack and will become the target of the self-defence action. It was on that basis that the US justified its self-defence action against Afghanistan following the '9/11' attacks and the same reasoning was offered more recently to justify the US action in Iraq and possible Syria against the Islamic State (ISIS/ISIL).

The immediate question is whether this construction can apply to the regime of state responsibility? According to international law, a state has a duty of due diligence not to allow 'its territory to be used for acts contrary to the rights of other states'. The duty of due diligence places an obligation on States to interfere with private actors and private conduct in order to streamline their behaviour in view of the State's international law obligations. Moreover, it creates an expectation as far as other States are concerned that the State acts diligently and respects its international obligations as well as the rights of other States. Since private actors undertake activities that belonged to States or in other words belonged to the public sphere and since private actors may be quite powerful to affect other States but the only subject of international law is still the State, the duty of due diligence underpins the international legal order. If a State breached its due diligence obligation, it will be held responsible for the breach but not for the ensuing act. Yet, the resultant act may be more pernicious, for example a destructive terrorist attack. What Professor Tsagourias proposes then is to combine the due diligence obligation with causation in order to provide a more effective regime of responsibility in the cyber context.

If this construction is applied to cyber, it means that a State that fails in its duty of due diligence to maintain cyber hygiene or to prevent wrongful acts being committed from its territory or infrastructure will be responsible for the wrongful act committed by a non-State actor if that act would not have happened had the State acted diligently. This is not a novel interpretation of the law of State responsibility. The same reasoning was followed in one of the founding cases of the law of State responsibility, the *Corfu Channel* case. In that case, although the involvement of Albanian agents in the mining of the Corfu Channel was not established, Albania had a general duty of due diligence which included an obligation to prevent any damage from the moment it learned about the mines. Because Albania failed to act, it was responsible not just for dereliction of its duty of due diligence but for the explosions and the damage and loss of life that resulted from them.

Professor Tsagourias concluded that as long as States are the primary actors in international law they should live up to high standards and should also be held responsible not only when they commit a wrongful act but also when they cause it by creating fertile conditions for such an act to happen. The proposed model of responsibility based on due diligence and causation will increase a State's vigilance and provide a more effective way of ensuring respect of international law by States and non-State actors.

Dr Marco Roscini continued the discussion by asking the question "how can you ascertain that a State is responsible for a cyber-operation?" Obviously, evidentiary issues in relation to State Responsibility are not specific to cyber operations. Already we can see this from the Nicaragua case at the ICJ, which highlighted the problems in relation to covert operations in Central America. It is undeniable, however, that evidentiary problems are particularly significant in the cyber context because trying to understand who is behind a certain cyber operation presents significant technical problems.

Taking a look at the most famous instances of cyber operations allegedly conducted by a State against another State, we can see that there are really no more than suspicions and allegations on who is behind these attacks. These suspicions and allegations are based on, for example, the political context in which they occurred, whether they were particularly sophisticated, the nature of the attack etc. But this is nothing other than circumstantial evidence and leaves us very far away from evidence on who is responsible for the attacks.

In the context of cyber operations, evidence is needed, at least as far as attribution (the subjective element of an internationally wrongful act), on at least three levels: first, providing evidence of what computer/server/IP address the cyber-attack originated from. Second, providing evidence of the identity of the individual behind the operation. Third, evidence in relation to attribution - evidence of the link between the individual and the State to prove that the State is responsible for the cyber operation. In addition to evidence for the subjective element, evidence for the objective/material element of the internationally wrongful act is also required. For instance, it is still not known with sufficient certainty if Stuxnet caused damages to the Iranian centrifuges in Natanz and if so, to what degree. This is quite an important

question because it determines the level of the charge - whether it is a use of force or not - and subsequently the appropriate remedies.

The subject of Dr Roscini's discussion was to look at whether there are any rules on evidence that would apply to claims in inter-state judicial proceedings for remedies against damage caused by cyber operations. He looked at proceedings before the ICJ as there is no uniform body of the law of evidence in international law due to each Court/Tribunal functioning with its own rules. The ICJ, as the United Nations' main judicial organ is the organ with potential jurisdiction, with the consent of the litigants, to adjudicate on claims for State Responsibility for the violation of any primary rule of international law. The talk examined the evidentiary issues concerning cyber operations through looking at the 'burden of proof', the 'standard of proof' and, finally, the methods that can be used to establish proof.

Starting with the 'burden of proof', the litigant has the onus of providing the evidence necessary to prove a certain fact. It is normally the party that alleges a certain fact that has the burden to prove it by providing the necessary evidence according to the standard of proof. There are exceptions to this general principle, which is expressed by the Latin maxim "*Onus probandi incumbit actori*" (the burden of proving weighs on the plaintiff). They are that non-disputed facts and facts of public knowledge do not need to be proved. Certain commentators have argued, and some states as well, that there is a reversal of the burden of proof in the cyber context. This is motivated by the difficulties in identification and attribution of the responsible individuals or State. It has been argued that in the cyber context, the burden of proof should shift from the accuser to the State from whose cyber infrastructure the cyber-attack originated. It is this state that has to prove that it was not responsible for the cyber-attack or that it exercised due diligence to prevent the misuse of its cyber infrastructure by others to conduct the cyber operation. This point would be inconsistent with the jurisprudence of the ICJ on the matter. In the *Corfu Channel* case the ICJ clearly said that the exclusive control exercised by a State over its territory does not automatically entail responsibility for the wrongful acts occurring therein. The difficulties of discharging the burden of proof in this case, however, may allow a more liberal resort to inferences of fact and circumstantial evidence. Even beyond territorial control the Court has rejected that there could be a reversal of the burden of proof when the relevant evidence is in the hands of the other party, or when the case concerns an armed conflict.

Another issue in relation to the possible reversal of the burden of proof derives from the alleged application of the precautionary principle, which normally applies in international environmental law, to the cyber context. It has been argued that the application of this principle to cyber operations leads again to a reversal of the burden of proof from the accuser to the state from where the attack originated. The ICJ has expressed the view that the precautionary principle may be helpful in interpreting and applying a treaty but does not lead to a reversal of the burden of proof.

In light of the above, should a case concerning cyber operations reach the ICJ, it is unlikely that the Court would allow a reversal of the burden of proof just because of the uniqueness of the cyber scenario. This makes sense because reversing the burden of proof may not actually

yield expected results. For instance, in the DDoS attacks against Estonia, the botnets that participated unwillingly in the attack were from more than 100 countries. If you look at the fact that cyber operations were often routed through many States, again it would be quite difficult to ask all those States to provide evidence of the fact that they were not responsible for the operation or that they exercised their due diligence duties.

The fact, however, that it may be more difficult to discharge the burden of proof in the cyber context may affect the 'standard of proof'. The standard of proof is the quantity and the quality of the evidence that a litigant has to produce in order to meet the burden of proof. There is nothing in the ICJ Statute, or in the rules of the Court, that indicates what standard is expected by the Court from the litigants. The Court also, normally does not indicate the standard in the proceedings, and when it does it generally indicates it in the judgement, when is too late for the parties to take it into account in their arguments.

In addressing what standard of evidence would be required in a cyber case, there is obviously no case law to look at so one must look for indications in official state documents, strategies, official statements etc. and see how the State defined the standard of evidence.

For instance, the US Air Force Doctrine for Cyber Space Operations requires that attribution of cyber operations is proved with "sufficient confidence and verifiability." A report by an Italian Parliamentary committee requires that attribution is proved "unequivocally and with irrefutable digital evidence." Germany refers to "reliable attribution." A UK House of Lords document refers to "conclusive attribution." And a Dutch document requires, in the context of self defence against cyber-attacks, that the origin and the identity must be "sufficiently certain."

It can be tentatively interpreted that these references are equivalent to the "clear and convincing'" standard that is used by the Court normally, although not always consistently, in relation to State Responsibility claims. The clear and convincing evidence standard sits between the 'beyond reasonable doubt' standard - which is normally used in criminal trials - and the preponderance of evidence / 'balance of probability' standards. Again, this makes sense. The 'balance of probability' standards in the cyber context (and even more a *prima facie* one), because of the characteristics of cyberspace, may lead to suspicious claims of erroneous attribution. On the other hand, a 'beyond reasonable doubt' standard would certainly be unrealistic in the cyber context. That said, in one case the Court found that in claims for State Responsibility, when the charge is particularly serious - meaning when it involves the commission of international crimes - the standard should be higher than 'clear and convincing'. In the *Bosnian Genocide* case, the Court required that evidence be 'fully conclusive'. If this is applied in the cyber context, claims for reparations from cyber operations amounting to war crimes, or crimes against humanity, or acts of genocide will require fully conclusive evidence and not just clear and convincing evidence. The Court, in the same judgement, made a further specification. It distinguished between the evidence necessary to prove that a State committed acts of genocide and the evidence necessary to prove that a State did not exercise the due diligence to not prevent acts of genocide. In the latter case, the Court said that the standard of evidence is still stringent but not necessarily fully conclusive and so there would be a difference in the standard of proof required to demonstrate that a

State had conducted cyber operations amounting to international crimes and that required to demonstrate that the State did not exercise the necessary due diligence to stop its cyber infrastructure from being used by others in the commission of international crimes

In addressing the methods of proof there are a couple of interesting points. The Court usually prefers documentary evidence over oral evidence. It particularly favours official State documents and documents from international organisations. The problem with the cyber context is that often state documents on cyber issues are classified either in whole or in part because of security reasons. It is difficult to refuse to produce a document before the Court just because it has been classified for security reasons. However, there is no sanction. The Court may take formal note of the fact that a State has refused to produce the document but, in fact, in the two cases where this situation arose - the *Corfu Channel* case and the *Bosnian Genocide* case - the Court did not draw any negative inferences against the State that did not produce the classified documents.

In the cyber context you also have a lot of reports by (technical or not) think-tanks and NGOs. Yet, the Court does not view these documents very favourably. They may have a value, but only a corroborative one. They have a lower probative value than official state documents. This is even more so for press reports and media evidence, for example, the *New York Times* articles written by David Sanger that claimed that certain States - the US and Israel - were behind Stuxnet and conducted cyber operations against Iran. Press reports and media reports are treated by the Court with great caution. If the reports rely on only one source, or if they rely on an interested source, or if they do not specify their sources at all, the Court gives no probative value to them. When they have a higher standard of objectivity, the Court gives them a two-pronged probative value: they can be used to corroborate evidence provided by other direct sources (if they are fully consistent) and can contribute to demonstrate public knowledge of facts, which the Court may then take judicial notice of, but then this is as far as the Court goes with regard to non-official documents.

With regards to the problem of whether presumptions or inferences of fact may play any role in relation to the possible attribution of cyber incidents, the ICJ has again shown an increasing reluctance to draw inferences. When it does, it does so to protect state sovereignty. As mentioned above, following the non-production of documents the Court has not drawn any negative inferences against the litigant that has refused to produce the documents. In the context of the exclusive territory of control, in the *Corfu Channel* case, the Court has said that it cannot be concluded from the mere fact of the control exercised by a State over its territory that that State necessarily knew, or ought to have known, of any unlawful act perpetrated therein. It is only through other indications of State involvement that territorial control may contribute to prove knowledge. And somehow this is reflected in Rules 7 and 8 of the Tallinn Manual. If control of the cyber infrastructure is not sufficient in itself to prove direct responsibility or even knowledge of the attack, it can, however, have an impact on the methods of proof. The Court, in the same judgement, says that because of the exclusive territorial control over the area, the State is often unable to furnish direct proof of facts giving rise to responsibility. Therefore, that State should be allowed a more liberal recourse to inferences of fact and circumstantial evidence. Inferences, thus, may have a higher probative

value when the litigant is unable to provide direct proof of facts because they are under the exclusive territorial control of the other litigant. However, when proof is based on inferences these must leave more room for reasonable doubt and no inference can be drawn which is inconsistent with facts incontrovertibly established by the evidence. Of course, the Court will need to establish whether the state has 'exclusive territorial control' over the cyber infrastructure in question and this is linked to the ongoing debate on the creeping jurisdiction of states over the internet and cyberspace in general.

In relation to inadmissible evidence, there are no express rules on the inadmissibility of evidence in the ICJ Statute. However, it is obvious that evidence which is produced too late, or not in the prescribed form, are not acceptable. It is more interesting in the cyber context to see whether evidence collected in violation of international law is admissible as a proper method of proof or not. Cyber espionage may for instance be a useful tool to collect evidence of state responsibility for cyber operations but it has been argued that these activities are an internationally wrongful act. Indeed, cyber espionage, when it entails an unauthorised intrusion into the cyber infrastructure located in another state, is a violation of the sovereignty of that State. Assuming, and it is a big assumption, that these activities are inconsistent with international law, what is the probative value of the evidence so-collected? When the situation arose in the *Corfu Channel* case, the ICJ does not dismiss evidence that was unlawfully collected by the United Kingdom. Albania did not raise the issue, and the Court did not address the legality of the evidence, but rather observed that the purpose of collecting evidence to submit to judicial proceedings did not exclude the illegality of certain conduct. So, the fact that direct evidence is located in computers or networks of another State does not entitle the interested litigant to access them without authorisation. However, if the evidence is so-collected it may still be taken into account by the Court.

Dr Roscini, in concluding, stated that the burden of proof does not shift in the cyber context and that the standard of proof is not dissimilar from that applicable to other cases of State Responsibility. More probative value is given to the Court by official documents and with regard to non-official documents the Court may take them into account but only as secondary sources. Thus they may only be used to corroborate direct evidence and to establish public knowledge of certain facts. On the other hand, drawing inferences has been treated with great caution. Finally, the fact that evidence is obtained illegally – for instance, through cyber espionage - does not excuse the illegality of the conduct but also does not exclude the fact that the Court may take such evidence into account.

Dr Russell Buchan began by addressing the issue of recent media reports purporting that a 'cyber-Pearl Harbour' or a 'cyber-Armageddon' is imminent. Perhaps unsurprisingly, there has been a rush amongst international lawyers to determine whether a cyber-attack can constitute an armed attack under Article 51 of the UN Charter, thus permitting a State to use military force in self-defence. Given the high threshold ascribed to the concept of armed attack under international law, most cyber-attacks are more likely to constitute unlawful uses of force under Article 2(4) UN Charter or even more likely an unlawful intervention in State

sovereignty – indeed, Dr Buchan believes that we have yet to witness a 'cyber armed attack' within the meaning of Article 51.

In a decentralised legal system, States therefore look to the availability of (peaceful) countermeasures in order to respond to cyber-attacks – that is, a non-forcible otherwise internationally wrongful act but considered temporarily lawful because it is deployed with the intention of inducing a State to comply with its international legal obligations. In recent years academic literature has focused upon the circumstances in which a State can utilise countermeasures in response to State-sponsored cyber-attacks. However, the empirical reality is that the majority of cyber-attacks are committed by non-State actors. This immediately raises the question of whether and to what extent States can deploy countermeasures in response to cyber-attacks committed by non-State actors. This is the research question addressed in Dr Buchan's presentation. The question related to the deployment of countermeasures in response to cyber-attacks committed by non-State actors, not the deployment of countermeasures in response to non-State actors directly. This is because, at least historically but even in the contemporary era, international law is a State-centric system that does not permit countermeasures to be taken directly against non-State actors. However, because international law is a system based upon the sovereign equality of States, all States are under an obligation to not allow their territory to be used in a manner, or actors within their jurisdiction to act in a way, that is injurious to the legal rights and interests of other States. This is known as the doctrine of due diligence.

The doctrine of due diligence, being the corollary of the principle of the sovereign equality of States, has deep roots in international law, and can be traced back at least as far as the *Trail Smelter* case in 1941:

> 'under the principles of international law…no State has the right to use or permit the use of its territory in such a manner as to cause injury…in or to the territory of another of properties therein, when the case is of serious consequence.'

This principle – although initially developed in the context of transboundary environmental harm – has now developed into a general principle of customary international law; see for example the *Corfu Channel* case, the International Law Commission's (ILC) Draft Articles on the Prevention of Transboundary Harm (2001), *Pulp Mills* (2010), the *Seabed Dispute Chamber's Advisory Opinion* (2011) etc. We now see this principle being applied as an autonomous customary international law obligation in the context of terrorism, human rights abuses, and even international economic harm.

However, there is still some resistance to this obligation being applied in the context of cyberspace. There are two main reasons:

1) Some still regard cyberspace as a global commons that is *res communis*. The argument runs that States do not exercise territory or sovereignty in cyberspace, and thus no obligations of due diligence can be activated – actors are instead operating in a legal vacuum. While it can be argued that States do not exercise territory in

cyberspace, it is clear from State practice that States do exercise sovereignty in cyberspace, whether it be over actors or even information.

2) The ILC Draft Articles on Transboundary Harm refer to 'physical harm', which excludes 'socio-economic harm or similar harm'. Although of course cyber-attacks can produce real world physical harm, the ILC's understanding of harm would presumably exclude many cyber-attacks that produce harm which is confined to cyberspace, even though this harm in itself may be serious. However, wider international jurisprudence and indeed the *Tallinn Manual* both refer to negative effects manifesting serious consequences – which is clearly capable of encompassing serious cyber harm that does not produce physical harm.

If we are convinced then that States do owe a duty of due diligence to actors operating cyberspace, the next question related to the scope of this duty. Determining this is important because it will impact upon the character and severity of the countermeasures that can be justified if a breach of this duty can indeed be established. In the *Pulp Mills* decision the ICJ explained that the duty of due diligence is 'a duty of vigilance', 'a responsibility to ensure' that is placed upon States to implement regulatory and administrative structures to prevent or minimise transboundary harm. This applies to actors within its jurisdiction committing cyber-attacks but also to hijacked computers or servers that unintentionally transmit cyber-attacks.

Crucially, however, this is a duty of best efforts within the State's capacity to do all that is reasonable and appropriate in the circumstances; what would have any reasonable State in that position and with those means and capacity available to it have done? This 'best efforts' standard is important in the cyber context because States will inevitably have varying degrees of technological advancement. Thus, what is reasonable for one State may be very different from what is reasonable for another State. For example, some States may be required to adopt and implement legal rules prohibiting certain cyber activities, dedicate considerable resources to enforcing these legal rules, establish regulatory regimes to monitor and prevent cyber harm (such as CERTS). However, in relation to States that are less technologically advanced it may only be reasonable to expect them to notify affected States that transboundary cyber harm is underway or likely.

However, the duty imposed is not static. The content of 'due diligence' obligations may not easily be described in precise terms. 'Due diligence' is a variable concept – it may change over time as measures considered sufficiently diligent at a certain moment may become not diligent enough in light, for instance, of new scientific or technological knowledge. It may also change in relation to the risks involved in the activity. Thus, as States engage in knowledge transfer and capacity building – which occur dramatically and quite rapidly in terms of a State's cyber capabilities – there will be a correlative increase in the intensity of the duty they owe in terms of policing their territory and actors within their jurisdiction.

With regards to the knowledge needed by a State in order to activate the duty of due diligence, there are those that argue that a State must have actual knowledge of the injurious activity (*Corfu Channel* case: Albania 'must have known'; 'allow knowingly'). If this is the

correct standard, however, many States will avoid liability for cyber attacks committed by non-State actors by simply refusing to dedicate resources to monitoring and investigating the conduct of actors within its jurisdiction. Therefore, Dr Buchan would agree with the Tallinn Manual that the appropriate standard is one of constructive knowledge, where a State will owe a duty of due diligence where it 'should have reasonably known' in the circumstances. After all, this accords with the underlying rationale of the doctrine of due diligence which is to impose obligations upon a State to do all that is reasonable in the circumstances – and so in response to that which can be reasonably known – to prevent or minimise transboundary harm. But this does not impose strict liability upon States for all injurious cyber activities – given the anonymity and ubiquity of cyberspace it may be quite reasonable to determine that a State should not have known about the activity.

Where a State fails in its obligation to pay due diligence it commits an internationally wrongful act and is potentially subject to countermeasures by injured States. However, as countermeasures are themselves internationally wrongfully acts countermeasures are a risky business for international law, and for this reason are hedged with limitations and restrictions; notification, reversibility, preventing third party implications. Dr Buchan focused instead on the requirement that countermeasures must be 'proportionate', not least because the contours of this limitation remain unclear under international law and this is particularly so in relation to countermeasures taken in response to a State's failure to pay due diligence in cyberspace. Historically, countermeasures were considered proportionate where there was an element of reciprocity between the original internationally wrongful act constituting the countermeasure – an eye for an eye approach. However, such an approach is premised upon a theory of retribution and punishment. As a self-help mechanism to enforce international law countermeasures are designed not to punish or avenge but instead to induce a State into compliance with its legal obligations. The legitimate aim of countermeasures must therefore be to secure compliance with international law.

However, in pursuing this aim States do not possess an unbridled right to deploy massive countermeasures in order to 'shock and awe' non-conforming States into compliance. After all, it must be remembered that countermeasures are *prima facie* internationally wrongful acts and so their use must be restricted not just to prevent excessive damage to the State acting wrongfully but also to in the international legal order more generally. In consequence, determining whether the damage inflicted by a countermeasure is proportionate to achieving the legitimate aim of inducing compliance is therefore restricted and limited by reference to the damage sustained – 'Countermeasures must be commensurate with the injury suffered, taking into account the gravity of the internationally wrongful act and the rights in question' (Article 51 ICL ASR). So the nature and extent of the injury sustained is all important.

Determining the gravity of harm is of course important here. There is a significant difference between a State's failure to prevent cyber-attacks that disable key government communications and minor acts of cyber espionage or website defacing. Moreover, the issue of duration is also relevant here. The implications and effects of a cyber attack can become more deleterious as time passes and access to crucial online resources is being denied. In the case of access to online banking; one, two or even three days may not be particularly harmful

but as weeks and months pass the damage becomes far more severe. Such circumstances could justify a commensurate increase in the severity of countermeasures. Similarly, cyber attacks have the tendency to escalate as time passes. The DDoS attacks in Estonia for instance; botnets grow exponentially and more and more computers become corrupted and send requests for information. Viruses, too, also spread rapidly and have the tendency to become more pernicious over time. Where this occurs, there can also be an incremental intensification of countermeasures.

Finally, and importantly in the context of cyber, it should be noted that assessment of damage is not limited to the damage sustained by the victim State. As was explained in the *Air Services* litigation, and indeed reflected in Article 51 of the ILC's ASR, implications for broader principles and community interests can also be considered. In the *Air Services* litigation, for example, the damage sustained included not just the economic impact of France's internationally wrongful act against the US, but the impact of France's conduct upon international air travel policy more generally.

This could be potentially significant in the cyber context where, for example, a State fails to prevent actors within its jurisdiction from conducting acts of cyber terrorism – where the suppression of terrorism is clearly an important interest to the wider international community and particularly when this activity is conducted in a globally interconnected environment – such a failure of duty of due diligence would justify far more intensive countermeasures in order to push a State towards compliance with its due diligence obligations than, for example, acts of cyber espionage which seeks to obtain discrete economic information about a particular company within a particular State.

This Report was prepared by **Paul Stokes**, Intern in Public international Law at the British Institute of International and Comparative Law.

The Seminar was convened by **Kristin Hausler**, Associate Senior Research Fellow at the British Institute of International and Comparative Law (BIICL).