Creating Accountability for Global Cyber Norms

By James A. Lewis

The year 2021 saw all UN member states agree in the Open-Ended Working Group (OEWG) on a framework for responsible state behavior in cyberspace, based on norms developed in the United Nations Group of Governmental Experts (GGE) in 2015. The OEWG formalized global agreement on the 2015 norms. This consensus agreement means that the framework for responsible state behavior is now politically binding for all member states.

Global agreement on the obligations for responsible state behavior is a substantial step forward in building international cybersecurity in a rules-based environment. Unfortunately, international experience since 2015 has shown that agreement on norms, even when politically binding, is by itself not enough to ensure their observation or create stability in cyberspace. This has shifted discussion from what norms are needed to how to build accountability and what to do when norms are ignored. While the norms agreed to by all member states will ultimately reinforce international stability, to make progress, it will be necessary to develop a collective diplomatic strategy to improve the observation of norms and increase accountability when they are ignored. A strategy of sustained engagement and the imposition of consequences is necessary for norms to have effect.

Our assumption is that accountability for malicious cyber actions can only be strengthened if there are consequences for a state’s decision not to observe norms. If nothing else, a failure to take action in response to transgressions seems to only encourage opponents. An immediate task is to define the conditions for collective action. There has been an initial and informal agreement among like-minded democratic nations that accountability requires the imposition of consequences for a failure to observe norms, but several issues must be addressed. These include agreement on standards for attribution of the source of a malicious action and agreement on a proportional, lawful, and effective response. A collective approach is essential if efforts to create accountability are to succeed, and any response to a cyber incident will require political heft and sustained engagement.

---

1 This is a report from a larger project supported by the governments of Australia and Estonia. It reflects several roundtable discussions among diplomats from more than a dozen countries. The conclusions or any errors, however, are solely the responsibility of the author.
Applicable Norms and “Agency”

What norms are in fact ignored? Many of the OEWG norms focus on protecting critical infrastructure. However, cyberattacks on critical infrastructure are rare. In contrast, risks of international peace and stability from other kinds of cyber actions are already increasing. If the aim is to increase stability, the relevant norms are found in paragraphs 13(a) and (c), where states agree to prevent actions that threaten or are harmful to international peace and security and to not knowingly allow their territory to be used for wrongful acts. This latter norm is vital, as it makes clear that responsibility derives from sovereign jurisdiction (or a failure to exercise it). This makes the assignment of political responsibility to a state easier and clearer.

Equally important, the foundation for the OEWG norms is the binding agreement by all states to observe the UN Charter, where states have agreed to refrain from the use of force or threats to use force against the political independence of other states and have agreed to respect international law and fundamental human rights. These OEWG commitments and charter obligations are routinely ignored by some states that violate other nations’ sovereignty in ways that damage security, stability, and human rights.

One shortcoming with the UN-agreed norms is their lack of “agency,” or the capacity to take action. In theory, agency falls to the international community acting through the United Nations Security Council (UNSC), but in fact, this means the ability to reach agreement in the United Nations on a response to malicious cyber action is very limited. The United Nations is a consensus-based organization, with many members who have different and sometimes competing interests. Creating accountability has always been difficult for the international community, not just in cyberspace. Actions must be exceptionally egregious, and powerful states will put their national interests first. Additionally, states are understandably reluctant to agree to coercive measures—particularly those involving force or the threat of force—for cyber incidents where culpability is deniable, rules are still unclear, and damage (from a strategic perspective) is usually minimal. The absence of consequences for individual cyber actions makes it easy to ignore efforts to impose accountability.

Creating accountability is complicated for cyber actions, since almost all such actions stay below the threshold of the use of force, on which many rules are predicated. The international community is more likely to respond to the use of force, but few cyber incidents have had effects that approximate the use of force. Despite hyperbole, no one has ever died in a cyberattack, there have been very few incidents of physical destruction, and until 2020, cybersecurity had never come before the UNSC, making it a lesser issue in the eyes of many member states not directly involved in cyber conflict. This is changing as the risks of malicious cyber action have increased, as dependence on networks has grown, and as a result of increased great power competition. Even so, there is still some reluctance to act in response to malicious cyber action.
Diplomatic practice does offer an alternative approach that provides for agency. This alternative is more attainable in the near term. It requires a group of like-minded nations to first agree on their responsibility to observe cyber norms in their national practice. This requires a public commitment at senior political levels for these nations themselves to be guided by the OEWG framework for responsible state behavior. On the basis of this commitment, the second step is collectively responding to violations of norms. This collective response has rarely been undertaken for malicious cyber actions. Changing this is a crucial first step. Consistent action by a group of states to impose consequences for not observing agreed norms is essential to creating accountability.

An example might be the Western effort in the 1990s to persuade China not to sell missiles. This effort took several years of high-level engagement and a constant reiteration of the need for responsible states to observe the norms developed for such transfers (in the Missile Technology Control Regime). The issue was raised with China at every level, on every occasion, and during every topic of engagement. A talking point on proliferation was appended to every senior officials’ statement. Most importantly, it was not just one nation appearing to complain and suggest; it was done repeatedly and collectively every time Western leaders engaged with their Chinese counterparts. One of the most important steps that can be taken to increase observation of the 2021 UN norms is to make this kind of sustained engagement a routine expectation for state behavior rather than a sporadic objection when unspoken expectations for respecting sovereignty and law are flouted.

Raising specific instances of a failure to observe norms with the state responsible and asking for explanation and cessation is an essential first step. However, such efforts will be more effective when accompanied by persuasive evidence (and adequacy will be discussed later) and when carried out by many nations. This can be done in both private and public engagement. In the proliferation case, it was persuasive to share with China and others the most senior levels of intelligence (such as satellite imagery) that established responsibility. For cyber actions, there will be objections to such sharing on the grounds that this will inform opponents of U.S. capabilities, but the detailed indictments made by the Department of Justice—often accompanied by pictures of the individual government agent—show such sharing can be done.

This is different from “name-and-shame,” a useless strategy involving a public statement made safely in capital and without any follow-up action. These statements have no political weight. What is needed is a direct engagement with the leaders of culpable states and the credible threat of a punitive response. The key to a new approach is a consistent, collective, and direct approach at senior levels to object to a state’s failure to observe norms accepted by all responsible states. Every time ministers meet with a counterpart, they must raise the issue of malicious cyber action, using points developed in common with their partners and providing specifics of the act. The actions that have come closest to this have only occurred with the joint condemnations of China and Russia and the statements by President Biden and the EU high representative for foreign affairs and security policy, Josep Borrell. But if such statements are to be more than one-offs, filed away and ignored, they must be repeated by all other G7 nations and by other EU member states. Consistent and direct iteration on the need to observe norms must become standard diplomatic practice for like-minded states.

A comprehensive strategy to create accountability would define and articulate possible responses to reinforce an integrated diplomatic campaign. In developing a collective diplomatic strategy linked to

---

2 By “Western,” this report means democratic states rather than some geographic distinction. The term itself is an inheritance from the East-West divide of the Cold War, still appropriate in distinguishing democracies from the major authoritarian powers (who, ironically, remain the same as in the last century despite the outcome of the Cold War).
potential consequences, the issues of political attribution, the development of a menu of response actions consistent with international law and agreed norms, and proportionality in any response come to the fore as determinative issues.

**In developing a collective diplomatic strategy linked to potential consequences, the issues of political attribution, the development of a menu of response actions consistent with international law and agreed norms, and proportionality in any response come to the fore as determinative issues.**

**Attribution Is a Sovereign Responsibility**

Attribution is essential for creating accountability. In order to impose consequences (or create accountability), attribution is needed as justification to the international community and to convince political leaders of the need for action. This is not the attribution required by a court but information sufficient to persuade decisionmakers and both domestic and global audiences that a response to a malicious cyber action is justified. Political attribution creates the informational conditions needed to validate an action in response.

Interviews with officials in a range of countries show that attribution and proportionality remain sovereign decisions. The idea of a multilateral mechanism for attribution or some list of proportional penalties for misconduct would be unacceptable to most major powers. States will not be willing to relinquish the sovereign authority to decide the source of an attack and what form any response should take. No country will delegate this authority to a third party.

This means that a favorite academic proposal, for an independent, third-party entity to provide neutral attribution (see the attached bibliography), is very unlikely to win support in capitals. The desire to have an impartial body of experts examine a cyberattack and determine who is responsible is understandable, but these proposals underestimate the difficulty of creating such a process and overestimate the benefits of a process not connected to an entity capable of taking action. Discussions with numerous states show that they are not willing to turn over attribution—which they consider a sovereign responsibility—to a third party. They see it as a derogation of sovereignty. Political sensitivities and risks that attribution can create for a state are additional reasons that they are unwilling to entrust this responsibility to a third party. The larger and more powerful the state, the less it is willing to consider a private entity assuming what is ultimately a political function closely tied to foreign and security policy.

Discussions with potential third-party entities who have considered acting as a source of attribution are not reassuring. When asked directly what they would do if their investigation led to a major cyber power such as China or the United States, they demur on their willingness to attribute. They show a willingness to attribute when investigation leads to a few weak states, such as North Korea, or to cybercriminals (if it can be determined that they are not acting as a proxy for a state), but this reluctance to attribute actions to major cyber powers means that few if any cyber incidents of political significance would be considered by a non-state attribution organization.

Proposals for an international institution such as the International Atomic Energy Agency (IAEA) to conduct investigations into cyberattacks and identify their source face similar difficulties. The IAEA works
because it is based on treaty commitments and is linked to the UNSC. The IAEA derives its authority from the Treaty on the Nonproliferation of Nuclear Weapons (NPT). It is the NPT that gives the IAEA the mission and authority to investigate and to verify compliance with the treaty. The NPT provides an essential element of international legitimacy and authority for IAEA activities. Absent a similar treaty on cyber activities, a cyber verification entity would not have the same legitimacy and effectiveness as the IAEA. The IAEA has reliable technical standards for attribution based on the physical attributes of nuclear activity. Such standards do not exist for cyberspace.

The IAEA’s effectiveness is derived not only from its technical expertise but also from its relation to the UNSC, which UN member states have agreed has the authority to impose sanctions or authorize the use of force. This provides the IAEA with agency. The UNSC provides agency for taking action in response to noncompliance. The IAEA reports to the UNSC when a nation is failing to comply with its NPT commitments, and the UNSC can then decide what actions can be taken. Politically, it is strongly in the interests of nuclear weapons states to prevent nuclear proliferation in a way that preventing malicious cyber actions is not. This concern guarantees their support. It is unlikely that states possessing cyberattack capabilities would be willing to inspect themselves or be inspected absent some verifiable commitment by all nations regarding cyberattack. The absence of a treaty and the political commitment by states that would underpin a treaty makes an IAEA-like approach to cyberattack unproductive. Further, an inspection agency that lacked some organic linkage to the UNSC and its ability to impose penalties would also be ineffective. Therefore, the IAEA is not a relevant precedent.

Attribution of the source of a malicious cyber action will remain a national decision, and any construct for collective action must recognize this. A sovereign state has the right to decide who attacked it. States will not give up that right to an international attribution mechanism or some international arbiter for attribution, and such proposals are unlikely to ever find agreement. Attribution remains the prerogative of states, and taking collective action will depend on one state persuading others of the validity of its case.

**Political Attribution**

Norm 13(b) of the 2015 UN GGE report calls for states to be thorough in investigating an attack, and both the GGE and OEWG note that attribution is a complex undertaking and many factors should be considered. The obligation created by the OEWG-agreed norms is that “in case of ICT incidents, States should consider all relevant information, including the larger context of the event, the challenges of attribution in the ICT environment and the nature and extent of the consequences.” This was originally a Russian proposal made in 2015 and intended to lay the foundation for potential objections to any effort to attribute a cyber incident to a state. Of course, China and Russia (or most likely any cyber power) will deny culpability, but the goal is not to persuade them to accept culpability but to persuade national leaders and the global audience.

This prudent approach can help avoid misunderstanding, but the willingness and ability to attribute the source of an attack is critical for increasing accountability in cyberspace and giving meaning to internationally agreed norms. Attribution is ultimately a political act and is essential for creating accountability. Political attribution requires a political context. What are the requirements for political attribution? What are the factors that states should consider in attributing an attack? This list of factors provides an initial conceptual framework for states to consider as they use attribution as a tool to increase accountability and stability in cyberspace.

Attribution is not primarily a technical question. In the attribution of a cyber incident of political importance, the culpability of a state is being assessed, not of an individual. Criminal action can be left
to law enforcement and the rules used by courts to establish culpability, but these are inappropriate for international relations. Accountability requires a state or a group of states to confront another state for its behavior. There is no court that can do this since the state actors in question will not only ignore any opinion from a court (especially as a court may lack the ability to enforce its opinion) but also question its legitimacy in declaring such an opinion. This changes the nature of the information required for attribution and response.

**Factors for Consideration for Political Attribution**

In forming a conclusion on attribution, the state must consider the weight, credibility, and importance of the aggregate information to determine if a preponderance of the information supports attribution to a particular state. After a review of significant cyber incidents stretching back more than a decade, we can conclude that cyber actions are not random but are taken in support of state policy. Since state cyber actions are guided by national strategies and foreign policies, these provide an indicator of intent and can be suggestive of culpability. The following 12 criteria present questions a victim state could ask to establish attribution and responsibility.

1. Who is likely to violate our sovereignty, judging by their public actions and statements?
2. Whose strategic interests are best served by violating sovereignty through cyber actions?
3. Do the effects of this violation support the strategic interests of another state?
4. Does the target of the attack (e.g., political or military advantage) point to state action?
5. Does the effect of the cyber action (e.g., what data was exfiltrated, what services were disrupted) suggest a state or state proxy actor? (Criminals are unlikely to go after military targets, for example, while states, other than North Korea, are unlikely to engage in financial crime.)

These can often be determined from publicly available information, national assessments, and consultations with allies and partners. Unfortunately, as international relations continue to deteriorate, certain states have become more open in expressing confrontational intent, making this kind of analysis of their policies easier. While in a court of law these factors might be viewed as circumstantial, in international relations a statement of hostility by a national leader is not made lightly, and patterns of action are indicative and sufficient to justify a response.

6. Do previous incidents, either here or in other countries, point to a particular state as the responsible actor? Are there precedential actions by the suspected state actor?

Cyber actions by states are undertaken in support of larger foreign, military, and technological policies. This creates a degree of consistency and predictability in who they select as targets, what information they extract, or what effect they attempt to produce. In many instances, more than one Western country is the victim of a larger campaign. The patterns indicated by this consistent behavior stretch across individual targets and can be an indicator of who is responsible.

7. Have other states experienced similar violations, perhaps simultaneously, and have they attributed the source of the violation?
8. Have allies or partner nations supplied supporting information on the source of an attack?
9. Does this incident appear to be part of a larger campaign?
Given that state cyber actions are intended to support larger strategic goals (such as undermining democracy, damaging NATO or other alliances, or gaining technological or commercial advantage), friendly nations are likely to experience similar actions in violation of their sovereignties. Even if they have not attributed the source of the action, their experience can indicate a pattern that can buttress conclusions on attribution. Many cases involve attributing one incident in a linked chain of incidents—in other words, a campaign. In other cases, a single cyber operation uses the same techniques against multiple targets in several different countries. These reinforce the ability to make a political attribution.

10. Are there technical indicators or other intelligence data that point to a perpetrator (these could be from a national source, a commercial firm, or an ally or partner)?

11. Is there supporting information from private sector actors, particularly cybersecurity companies, or supporting intelligence from national sources and from allies and partners?

12. Have relevant national services or these external sources been accurate in past assessments of attribution?

Although the needed information may not be immediately available, intelligence reporting can strengthen conclusions on the source of an attack. This should go beyond similarities in techniques and tradecraft, since by themselves these are not conclusive, as they can be copied and spoofed. However, on combination with other factors, such as indicators and analysis of attacker strategic intent, similarities to other incidents increase the probability that an assessment of culpability is correct.

Many cyber attribution successes are the result of successful counteroperations. Some are the result of the “exploitation of the attackers’ operational security errors.” This sort of evidence is likely to be more persuasive with political leaders, who routinely demand a high degree of confidence in any instances where they are asked to contemplate countermeasures. While cyber intelligence can be the most likely source of confirming information, traditional intelligence from agents in the attacking country or from signals intelligence can also buttress any conclusion.

Effect need not be determinative for attribution. There is of course a greater imperative to attribute the source of a damaging attack. But if the goal is to change behavior, it is not the success or failure of an individual action that is determinative but rather the decision by the attacking state to violate the sovereignty of the victim and ignore agreed norms that is important.

It is essential to recognize that political attribution involves assessing the culpability of a state, not an individual. This is not a criminal proceeding. Political attribution does not require identifying the specific individual(s) responsible for an attack. While this individual identification may be beneficial, the most important thing for political attribution is identifying the state responsible for the attack (or for failing to observe its obligations under agreed norms and international law to prevent its territory from being used for malicious action).

It is essential to recognize that political attribution involves assessing the culpability of a state, not an individual.

Practice in the physical world has created a predilection for catching the perpetrator red-handed. Cyberspace is a different environment and different rules apply. Absolute certainty can be a scarce commodity in the world of covert action where most cyber incidents occur. There will be some instances
where intelligence information provides a high degree of certainty, as with cases involving Russia, China, and North Korea that included images of the actual perpetrators, but these are rare relative to the number of state-sponsored cyber incidents and have often been the result of counter hacking the attacker’s network or receiving corroborating information from other, non-cyber intelligence sources.

Other factors help determine the degree of rigor required for attribution. The most important of these factors is whether an attribution is tied to the imposition of consequences against the attacker and whether the conclusion of an effort at attribution will remain internal or will be made public. An internal assessment creates little diplomatic risk.

Attribution will be subjective, and while political leaders will demand a high level of confidence (particularly if they are to impose consequences), the evidentiary standard cannot be complete certainty but sufficient certainty to justify action (noting that the first actions will be to consult with allies and partners for supporting evidence and to inform the offending state of initial conclusions).

While false-flag operations are occasionally used by attackers, they are uncommon. The most successful involved Russian activity against France’s Canal 5, first attributed to the Islamic State but quickly corrected by the French to identify Russia. It is possible that attackers overestimate the likelihood (and hence the risk) of being detected. It is also possible that larger state opponents no longer care if they are detected (according to some U.S. officials). The point of imposing consequences is to change this sense of invulnerability.

Charges or misattribution from the accused state or assertions that attribution is part of some larger and nefarious Western plot involving espionage and interference are part of the normal repertoire of authoritarian diplomacy. Despite this, intentional misattribution of a specific incident is very uncommon. This may be because consequences of misattribution are significant for the credibility of the accusing state. State practice and states’ desire to manage risk—in this instance and in others—differ significantly from hypothetical scenarios of misattribution. Given the increasing number and severity of incidents, as well as the growing lack of concern over being detected by leading opponents, it is not a major issue. There are so many incidents that ones with uncertainty can be replaced—the context being two decades of attacks and hundreds of incidents.

**Collaborative Attribution**

Attribution in the context of accountability is primarily political and requires a sound basis of intelligence for political decisions. However, the current level of information sharing among states is inadequate. While attribution is not primarily a matter of technical capacity, a lack of capacity is an impediment for many states. Additionally, there will always be some desire for independent confirmation or, when that is not possible, for independently drawing conclusions from the available information.

Many states are working toward developing their attribution abilities, but some do not yet have adequate capabilities. While attribution capabilities can vary widely among countries, private sector actors, such as FireEye or Cloudstrike, now have the capability to attribute the source of a hostile cyber action. This progress in attribution is not recognized in the international community, and there is a lack of agreement on what level of attribution is required for cooperative action among states. However, these issues can be remedied and should not prevent steps to increase accountability or impose consequences. Creating a framework of technical and factual attribution, combined with the political decision to act, could help frame the matter and draw attention to the factors required for action. There are many different techniques for technical attribution. Each has strengths and weaknesses, and no single technique replaces all others. Ultimately, technical attribution is not a substitute for political decision.
In the past, intelligence-based attribution has faced obstacles in sharing with public or foreign audiences. Governments have been reluctant to share the technical or espionage details of an attribution given the sensitive nature of the information. Sharing is also complicated by the disparate technical capacity among partner nations, but the development of common understandings and practice at regular sharing can create a conducive environment where trust is greater and sharing can be expanded.

States can use multilateral, regional, bilateral, and multistakeholder platforms to exchange cyber attribution practices and share information on the attribution of different types of ICT threats and incidents. These include developing common evidentiary standards and information-sharing mechanisms for coordination of collective attribution. Coordinated attribution of malicious activity will require better information sharing between partners, and perhaps new mechanisms for sharing and harmonization, but will greatly strengthen the political effect of any accusation.

**States can use multilateral, regional, bilateral, and multistakeholder platforms to exchange cyber attribution practices and share information on the attribution of different types of ICT threats and incidents. These include developing common evidentiary standards and information-sharing mechanisms for coordination of collective attribution.**

**Creating Accountability**

Attribution is a first step but is not by itself sufficient to change behavior. Decisions on attribution and on response are different. A state may well know with sufficient confidence who is responsible but choose not to respond. The issue now is what level of certainty is required to justify a response to an incident. Deciding how to act upon evidence requires a political decision by those ultimately responsible for the nation’s interest.

The community of norms-observing states can create accountability by using the 2015 GGE norms as a framework for action. The goal is to reduce the number, scope, and risk of malicious cyber actions. This will require mechanisms for cooperation and common understandings of attribution, proportionality, and risk management for responsive action. The ultimate arbiters of responsible behavior are other states, operating in the political process of diplomacy.

Devising ways in which states can collectively impose consequences (such as retorsion or more intrusive and coercive countermeasures) will be necessary to increase accountability. This raises the issues of what kinds of consequences should be imposed for malicious action and how this can be done collectively. States will need a broadly accepted menu of possible consequences and an ability to ensure that any consequences imposed are both proportional to the initial incident and consistent with international law and practice. This latter point is essential for winning international support for (or at least acceptance of) accountability measures.

The best approach to improving attribution and accountability is voluntary collective action by like-minded states. This needs to be considered in relation to existing measures (such as the Framework for a Joint EU Diplomatic Response to Malicious Cyber Activities), the obligations of international law, and the implications of other treaty obligations for collective self-defense.
We can look to diplomatic precedent for the means to impose accountability. Strategies for accountability must reflect long-established state practice on the use of diplomacy and force but go beyond the “name-and-shame” efforts of both private and public actors. When norms are violated, the initial response is diplomatic. Individual and collective engagement with the transgressor state aims to obtain agreement to stop the offending action, accept responsibility, and agree to avoid such actions in the future. For example, the efforts to persuade China to observe nonproliferation norms involved many nations engaging at senior levels and repeatedly raising the issue in every meeting with Chinese officials.

This kind of “full-court press” has never existed for cybersecurity. While China is now in a more powerful and assertive position, like-minded states will need to resort to the kind of coordinated diplomatic pressure used in the past, potentially accompanied by various punitive economic and diplomatic measures. Similar coordinated action is needed for Russia. For powerful states, this coordinated action entails joint condemnation by the international community and collective action on possible punitive measures (noting that many states feel that sanctions alone are not sufficient to change behavior).

One task is to define and agree on the conditions for collective action. In a September 2019 joint statement on advancing responsible state behavior in cyberspace, 28 nations agreed to “work together on a voluntary basis to hold states accountable when they act contrary to this framework, including by taking measures that are transparent and consistent with international law.” This, along with the norms endorsed by the OEWG in 2021, provides a foundation for expanding the observance of norms and increasing cyber stability.

**The Need for Accountability**

Consequences entail the range of internationally lawful responses available to states who are the victim or target of malicious cyber actions that run counter to agreed norms. While there has been some progress toward the imposition of consequences, such as sanctions by the European Union or indictments and sanctions by the United States, this has so far been an ad hoc and episodic process that has not increased global observation of norms nor the degree of stability in cyberspace.

The traditional levers of international relations are based on economic, political, and coercive measures tailored to the gravity of the incident, the capacity of the victim state, and the nature of the offender. For example, Russia is less likely to be affected by trade measures than China but more likely to be affected by actions affecting its international financial position. China is a more difficult target given the deep commercial connections that have developed since the 1980s. China’s economic strengths will, at least initially, make it more difficult to get consensus among Western states on how to respond.

The political concerns that shape consequences are similar to the political concerns that shape attribution. Credible attribution is essential for a politically acceptable response to malicious cyber action—acceptable to domestic publics, to allies, and to the global community. Similarly, any response based on that attribution must be seen by the same audiences as proportional. Proportionality is complicated because responding to a single incident in equivalent terms will dilute the effect. Responses cannot be sporadic, reactive actions if they are to affect opponent behavior. This is not the proportionality described in the Geneva Convention and requires a different calculus to gauge risk and political effect. Efforts to avoid all risk would be paralyzing and only encourage continued attacks. This

---

3 “An attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated,” Article 51(5)(b) of the 1977 Additional Protocol I.
calculus will require experience and a careful estimate of what actions will be seen as both justifiable and effective in leading the target state to recalculate the risks of continued malicious actions.

Political leaders’ desire for certainty in attribution before taking any action reflects their concerns over the potential risk of any response and their desire to avoid unintended consequences that could inflict harm on civilians. The first risk is somewhat overstated—there has never been an incident of escalation in the 30-year history of cyberattacks, and escalation risk is manageable. The second risk poses a more complicated problem. Broad economic sanctions, by their very nature, harm civilian populations. Targeted sanctions on individuals or agencies appear only to annoy our opponents and are not sufficient to change behavior. But one unspoken rule in cyber conflict is staying below the threshold of the use of force (generally taken to mean physical damage or casualties).

While a few Western nations have begun to explore the use of offensive cyber operations to create consequences, these have been few in nature, focused on private actors (possibly with the exception of the UK response in the Skripal incident), and have occurred without noticeable effect. One reason for the lack of effect may be that any response has been episodic—a one-off. This points to one central issue for creating accountability: Is this a response to an individual action, as would be the case in a law enforcement model, or is it a response to a sustained opponent campaign?

Incident or Campaign?

Attribution is just a tool for pursuing the larger goal of creating accountability, not an end in itself. One key issue—similar to the issues raised in the determination of proportional response—is whether to act against a single incident or a pattern of incidents. Cyber actions almost never involve the use of force or entail consequences similar to incidents involving force, but they do involve the violation of sovereignty. The routine and increasingly flagrant violation of sovereignty for more than two decades without coherent or direct objection is one reason for the parlous state of international cybersecurity.

In this, it is important to break the particularistic jurisprudential mindset of tit for tat. The response to each incident cannot be limited to an equivalent response. Years of sustained malicious cyber action by Russia and China amount to a campaign, not a series of sequential but unrelated incidents. This means there is an aggregate of malicious action. Political attribution does not require recognizing the attacker in every instance, but it does require identifying those incidents that can be used to support the objectives of a diplomatic campaign to change the behavior of China and Russia and win international support for doing so.

The occasional objection to a SolarWinds or a Hafnium will not produce the desired effect. The incident itself is not the problem. The problem is the consistent use of such actions by two major states. At this time, after more than two decades of hostile cyber action, there are undeniable patterns. Ideally, political attribution would combine public statements of antagonistic intent from the attacker, a pattern of hostile cyber activity that correlates with the attacker’s strategic objectives, corroborating information from allied and partner nations, some technical intelligence, and confidence in the analytic capabilities of one’s own services. This “package” provides reasonable certainty.
this time, after more than two decades of hostile cyber action, there are undeniable patterns.

The definition of coercion in cyberspace is complicated and not fully agreed upon. At a minimum, it means threatening to inflict or inflicting damage (either tangible or intangible). Coercion includes nonforceful actions, such as sanctions or indictments. These also have effect but are unlikely to create meaningful consequences for a failure to observe norms, perhaps because they have so far been imposed in a sporadic and uncoordinated fashion. For example, sanctions are ineffective against Russia, which has become inured to them. Sanctions are also generally not sufficiently harmful to affect China. Indictments, while painful, are too narrow in their effect. Anything short of a more forceful response than what has been seen to date (and this can be diplomatic rather than coercive) is likely to go unnoticed by opponents who expect coercion and threats to be a normal part of international relations. Opponents should come to expect to be routinely called out and penalized in some way in response to malicious action.

Proportionality and Risk

Response actions must be proportional to opponent action both to meet obligations under international law and to manage risk of expanded conflict. But proportionality remains unclear in cyberspace. What is the proportional response for election interference? Cyber actions do not map well to conventional concepts of force and coercion. Defining proportionality may require a period of experimentation, since cyber actions blur the lines between military, espionage, and political warfare in ways that require adjustments to thinking, doctrine, and tactics.

Many analysts incorrectly conclude that responsive cyber action must entail attacks on critical infrastructure. This would be unwise in the absence of some larger armed conflict. Cyberattacks on critical infrastructure by or against a nuclear armed state are exceptionally risky given the potential responses, making them both ill-advised and unlikely. The most effective targets are political and informational, but choosing such targets complicates the issue of proportionality.

It is important to note that consequences are very unlikely to involve military action and will stay below the use of force threshold implicitly observed in cyber conflict. Militarizing consequences would be counterproductive, if only by making it difficult or impossible to win broad international support from allies, partners, and a global audience and by raising issues of proportionality. There are many other punitive or coercive actions available that can be used to create accountability.

A precedent for a menu of consequences can be derived from the use of cyber operations in actions against the Islamic State. These operations, conducted by Task Force Ares, used cyber tools to degrade the Islamic State’s command and control and internal messaging, financial support and fundraising network, online recruitment and training, propaganda, and public communications. Similar targets can be found among cyber opponents. Temporary interference with command and control was undertaken against the Russia Internet Research Agency in 2020. Such consequences send a useful message, especially when not a one-off event. Other precedentual actions include the leaks of the Panama Papers, Chinese official corruption, and the publication of details of Putin’s new multimillion-dollar dacha. These were effective because they engender a vigorous, even angry, response. Corruption is a better target than critical infrastructure.

Exploiting opponent political structures could be usefully coercive. This does not mean attempting to tamper with their voting machinery, which would be a waste of time and effort, but exposing the mechanisms of influence and control used by Russian or Chinese political elites, along with their internal
conflicts. It could also mean information operations to expand discontent—possibly being particularly effective against China. But such actions would need to be carefully calibrated, since efforts to undermine their regimes are what U.S. opponents fear most, potentially triggering a more aggressive response.

For example, when a U.S. media outlet published the details of the wealth illegally assembled by a Chinese leader’s family, there was a prompt, aggressive, and covert response by Chinese intelligence agencies as they sought to find the source of the leaks, identify links to the U.S. government, and punish the media outlet. The fragility of opponent regimes, whose greatest fear is their own populations, means political actions will produce a reaction greater than what is needed for coercive purposes.

Political interference could be used in small doses to make the point, but a sustained campaign to undermine opponent governments, as Russia has done to the United States, would only make sense in the context of a decision about what the United States wants in its relationship with Russia. Is the goal to replace current leaders or to use specific actions as cautionary examples to dissuade them? An effort to create regime change faces many difficulties, may provoke a violent response, and might lead to a replacement worse than what exists now. In any case, regime change in Russia, China, or other hostile states is not Western policy. Before using political action, the United States needs to decide what it reasonably expects to achieve with those countries and how best to use measured coercive acts to change behavior.

While a response menu must be graduated by effect, the risk is that the lower actions may not be sufficient to change opponent behavior. Authoritarian states have in many instances a higher tolerance for pain (at least in regard to their citizens) than democracies. This is suggestive of what pressure points are most likely to be effective, but it also requires a different perspective for responses, since those that may be most appealing to democracies may be insufficiently effective. The matter is complicated because it may take many interactions and some time for response actions, even if repeated, to affect opponent behavior. An initial approach could use an iterative process that combines sanctions, public condemnation, and diplomatic actions to be repeated each time there is a violation of agreed norms.

Any responsive action must fit into a larger diplomatic strategy of cooperation and engagement. Actions by one nation, no matter how powerful, are likely to be ineffective. The same is true for responses that are sporadic and occasional. A first step could involve private discussion with the malicious actor to explicitly warn in advance of certain retaliatory actions if malicious cyberattacks continue. This kind of warning is probably essential. There is of course a credibility problem for Western governments after years of inaction. A second step would be to carry out the action and publicly make clear the grounds for doing so, including detailing attribution and explaining why the response is deemed proportional. A third step would be to warn the opponent to expect further action if there is no change in behavior and to consider whether to bring the matter to the UNSC—political escalation that raises the stakes for continued malicious action. The basis for doing so would be the inherent right of self-defense found in Article 54 of the UN Charter. Involving the United Nations will increase credibility with other nations and make clear how any action is consistent with international law.

The menu of responses is less important than an opponent’s belief that it will be used. To change behavior, responses must be part of a larger and coordinated diplomatic campaign that seeks to build international support and engages opponents at senior levels on a consistent and regular basis, with the ultimate goal of bringing opponents to the negotiating table to agree on concrete measures for the observation of norms. The chief variable is the degree to which Western nations believe they can reach an accommodation and the degree of commercial, political, or security risk they perceive in taking action in response. Efforts to
expand accountability must be part of a growing consensus among democracies on the need to respond to the challenges posed by hostile authoritarian regimes.

Our goal must be to reach some understanding with authoritarian nations on what is no longer acceptable in cyberspace. Imposing consequences to improve one's negotiating position is a long-standing diplomatic practice and an achievable objective for cooperation. Meaningful negotiation with U.S. opponents is not possible in the near term—they are not interested in serious engagement and the United States has a credibility deficit—but if the United States pursues an assertive strategy, it can eventually bring them to the table. In any negotiation, the West will not get everything it wants. China and Russia will not observe human rights or the rule of law and regard these as unjust violations of their sovereignty. Pursuit of these objectives must use other techniques. But the United States can improve its situation in cyberspace by building on the 2021 norms and creating accountability with the regular imposition of consequences for failure to observe them.

This discussion has laid out possible new directions but leaves many open questions. Some can only be answered in light of experience. There is risk in this, but risk is unavoidable if change is desired, and risk can be managed. The ultimate test for policy and strategy is actual engagement. What form consequences will take, what messages should accompany them, and how much time is needed for opponents to react and adjust in light of a new diplomatic approach are all open questions. For norms to have effect, actions need to be identified that create political effect, are consistent with obligations under international law, coordinate among allies, and communicate intent to publics and opponents.

James A. Lewis is a senior vice president and director of the Strategic Technologies Program at the Center for Strategic and International Studies in Washington, D.C.

This report is made possible by support from the Estonian Ministry of Foreign Affairs and the Australian Department of Foreign Affairs and Trade.

This report is produced by the Center for Strategic and International Studies (CSIS), a private, tax-exempt institution focusing on international public policy issues. Its research is nonpartisan and nonproprietary. CSIS does not take specific policy positions. Accordingly, all views, positions, and conclusions expressed in this publication should be understood to be solely those of the author(s).

© 2022 by the Center for Strategic and International Studies. All rights reserved.
Select Publications on Cyber Attribution: 2017–2021


