Journal of Military and Strategic Studies

VOLUME 12, ISSUE 2, WINTER 2010

The Necessity of (Some) Certainty - A Critical Remark Concerning Matthew Sklerov's Concept of "Active Defense"

Dr. Sandro Gaycken – University of Stuttgart

Cyber deterrence is clearly important. Secure passive cyber defense is impossible, so deterrence is the only feasible path. However, deterrence is pointless without attribution. This is logical from a strategic point of view. If retaliation does not hit the attacker, he will not be deterred. And it is of legal importance as well. Retaliation against the wrong actor is unjust and a crime of war. Thus attribution is a necessary condition for the law of war. An attacker has to be identified and, to make it an armed attack and not just a criminal act, the attacker has to be a state actor.

This last problem, the attribution of agency, is what Matthew Sklerov addresses centrally in his recent paper, *Solving the Dilemma of State Response to Cyberattacks: A Justification for the Use of Active Defenses against States Who Neglect Their Duty to Prevent*.[1] The problem stems from an actual and pressing situation. Hackers from Russia or China attack other countries in what could be considered armed attacks. However, the attacks have never been attributed to these states as Russia and China denied any involvement claiming that private hacker groups, not under their control, were responsible. Many observers now share the opinion that this is only half-true. At least, both countries seem to tolerate the attacks on foreign, non-allied states. Yet it is impossible to prove. Without proof, it is not possible for the victim states to retaliate and deter future actions against them. As these attackers are officially criminals, the law of war forbids deterrence. Sklerov calls this the "response crisis". To solve this problem, he develops an intriguing

[1] M. Sklerov, Solving the Dilemma of State Response to Cyberattacks: A Justification for the Use of Active Defenses against States Who Neglect Their Duty to Prevent. In: 201 MIL. L. REV. (1-85) 2009.

ISSN : 1488-559X

argument. He suggests that states have an obligation to prevent non-state actors, acting from within their states, from committing armed attacks to the best of their abilities. If they do not comply with this obligation, it is legitimate to attack them in anticipatory self-defense. This is what he calls "Active Defense". By this measure, states who tolerate attacks by non-state actors could be forced to end their tolerance and these attacks would subside.

This part of Sklerov's argument is sound. It stretches some interpretations of anticipatory self-defense and state responsibility for the actions of non-state actors to a degree some people might not feel comfortable with. While that is debatable, what is less debatable is Sklerov's idea of the determinacy of the identity of the attackers, the other necessary element of attribution. Not only has the type of actor need to be identified but the location needs to be identified as well. Following Sklerov, this element of attribution is indispensible. Accurate identification of an attacker's location is a clear necessity to support his argument. A state might be held responsible for armed attacks carried out by its own citizens, but it can hardly be responsible for armed attacks carried out by citizens of another state.

To trace an attacker's location with certainty, however, is not as easy as Sklerov sketches it. The only thing a victim of a cyber attack can determine with certainty is the location of the computer which led the immediate, last strike against it. But cyber attacks can easily be relayed via a whole chain of computers. Cyber criminals, for instance, regularly use intermediary computers, hijacked previously to their attack. These intermediary computers are usually distributed worldwide, in a number of countries. This process is called *routing* or *server hopping* and, to a certain degree, it is a standard feature of the internet. Routing by use of a number of predetermined hijacked servers makes tracking extraordinarily difficult. This is due to the structure of the internet. Technical lay people might think that any attack leaves a clear path which can be traced back. But this is not the case. If a cyber attack (information) is sent via a number of servers, these servers only remember the last server which sent the package and the next server to receive the package. They do not remember the whole chain of addresses. What is even worse is that any information about the connection is usually destroyed after the interaction. This is necessary as connections are undertaken millions of times each day, every time someone connects to the web, and it would consume far

too much storage space. Service providers in some countries are forced to keep records about connections for a while, but these are still only a few.

The overall situation makes it extremely difficult to trace cyber attacks. If the attack is ongoing, an analyst could in principle hack his way back to the attacker. But he would have to be capable of hijacking the servers used for the routing himself which is very difficult and it needs to be undertaken in a very short time span. Most cyber attacks, however, are not detected while they are under way. A running attack will mostly look like a normal, authorized access. Network analysts will find that vast amounts of data have been transferred only after the attack. Many networks will never notice an attack. Thus, normally, the analyst must try to trace the attacker after the attack has taken place. This identification requires two things. First, the attack has to have been routed through countries which support the logging of connections. This is not likely as attackers know very well which countries log data and which do not. Cyber criminals are well-known to relay spamming or phishing attacks through African countries because they do not log connections. But, even if this problem could be solved, the full physical cooperation of any link in the chain would be needed as the log records cannot simply be accessed remotely. Remote access would require illegal hacking, the logs would be hard to find and they might even be stored on physically distinct servers due to data protection regulations. Administrative assistance of the routed country would clearly be needed. However, many countries are reluctant to provide this assistance for several reasons. Invited foreign analysts could abuse their access for espionage. Also, some data would need to be protected due to its sensitive and private nature. As well, national data protection laws will forbid the unlimited access needed to fully analyze an attack. Accordingly, neither the mostly private owners nor the states in which they live are very likely to cooperate. Finally, the address information can also be spoofed or anonymized. There are a number of tips and tricks to avoid being traced, so specialists need to make detailed and time-consuming forensic analyses.

In other words, the identification of an attacker's location is anything but certain. Sophisticated attackers can easily circumvent detection. Sklerov admits that this is a problem and that there are severe limitations for tracing attacks. However, he tries to

escape this fatal difficulty by relying on an analysis by Wheeler and Larsen. They list a number of best-practice approaches for tracing cyber attacks.[2] This study, however, draws on a number of difficult presumptions. Some of its ideas have to be considered outdated. Honeypots, sophisticated filtering and spoof prevention techniques, for instance, have proven to be only of limited effectiveness. Attackers are usually smart enough to by-pass these and, as of late, they continue to be a step ahead. Other suggestions involving the cooperation of other entities are not practicable. The securing of servers and routers to prevent their abuse as intermediate systems in a process of routing is an example. The techniques prescribed to make them robust involve the removal of unnecessary or insecure services. But servers and routers are mostly held by private businesses, which earn their money with the provision of as many services as possible. There will be no incentive to remove services. Other suggested techniques are rather expensive in themselves. A continuous high-end vulnerability scanning and patching process for instance has to be bought from IT-security companies at considerable costs. Also, to be maximally secure, the software and even some parts of the hardware need to be state of the art. The best practice suggests updates in terms of hours. This, too, is expensive. Thus, many service providers will not be able to comply with these demands. The gravest problem, however, is that all demands would have to be enforced globally. The globality of the internet allows intermediate attacks from any place on the planet. So every host or server on the planet needs to be secured in this highly complicated and expensive way. Otherwise, an attacker will simply route his attack through those countries too poor or unwilling to cooperate. Some industry-led initiatives exist, proposing free updates, free antivirus programs and vulnerability scanning for developing countries as these are notorious, yet innocent hosts of (thus far criminal) cyber attacks. But it is unlikely that these initiatives will be successful to a satisfactory degree.

Other ideas put forward by Wheeler and Larsen do not comply with domestic and international law. One suggestion in immediate conflict with privacy and data protection laws is the surveillance of attackers. The reason for that is that the observers cannot know who might be an attacker in advance, thus leading to the logical

[2] D. Wheeler & G. Larsen, Techniques for Cyber Attack Attribution. In: INST. DEF. ANALYSIS, Oct. 2003, at 23–24, available at http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA468859&Location=U2&doc=GetTRDoc.pdf.

conclusion that everyone has to be put under surveillance. This is a major concern in current legal debates addressing illegal file-sharing. Such large-scale surveillance, however, has to be considered with great moral scrutiny. It implements a principle of "guilty unless proven innocent", in direct opposition to the presumption of innocence. A similar observation goes for the placement of sensors. They have different design approaches. Some of these approaches are in conflict with privacy concerns. Also, sensors should be placed close to an attacker, not close to a defender. This suggests yet again that the full cooperation of all countries of the world would be required to make the idea feasible. Another rather controversial, yet central idea is the implementation of analytical tools on foreign hosts and servers under control of US authorities. This conflicts with international law. It is a clear breach of the informational sovereignty of a country. It allows for the supervision of any kind of informational behavior of that country's citizens (sometime even of a number of countries). Additional difficulties will arise through conflicts with domestic data protection laws.

Finally, Wheeler and Larsen have to admit themselves that even their highly controversial method of tracking attacks is "inherently limited".[3] Delayed attacks cannot be traced. And many attacks currently are delayed. Successful attribution of culpability might simply fail in many cases. Finally, it might "identify the wrong location or identity of an attacker".[4] This problem could even be caused intentionally by an attacker. He could either disguise himself or provoke an attack on the falsely attributed attacker. The latter option also disables different approaches to attribution which want to infer attribution from the political context of an attack.[5] These approaches are not practical for a number of reasons. Determining a "political context" is very much open to interpretation. If attacks can intentionally be routed through actors with political contexts, this pattern of attribution becomes even more questionable. It is very likely to lead to false accusations while the actual attackers escape. Wars could easily be initiated by third parties. Owens, Dam and Lin termed this

[3] Sklerov, p. 51.
[4] Ibid.
[5] R.L. Kugler, Deterrence of Cyber Attacks. In: Kramer/Starr/Wentz (Eds.) *Cyberpower and National Security*, Dulles 2009

"catalytic cyber-conflict"[6]. No quantitative data exists to date for any of these potential failures tracking the instigators; there is no documentation recording the number of successful traces compared to failed ones. Security professionals share the impression that the ratio does not favour ascription.

In sum, the Wheeler and Larson paper is not a good point of reference. And the past five years have not changed this situation significantly, despite the existence of a few new tricks and services such as darknet monitoring or blacklist services. The problem of connecting attack and attacker is notorious. It is part of the very structure of the internet. Notwithstanding these concerns, Sklerov claims that "automated or administrator-operated trace programs can trace attacks back to their point of origin".[7] This sounds strange. IT-security professionals doubt that anything like this could exist.[8] Many of the brightest in the industry repeatedly tried to come up with trace programs, but were unsuccessful. Only less serious companies claim to have actual solutions. Any existing technologies will be immature, imprecise and quite likely in conflict with domestic and international law. This severely restricts Sklerov's approach. Even if the attribution of the type of actor can be allowed to be imprecise, the attribution of the location cannot. If there is a likelihood of, perhaps, 50 percent that the assumption about the location of an attacker is plain wrong, is that considered sufficient reason for an armed attack in anticipatory self-defense?

The conclusion is, that, despite the fact that he has a well-argued case for the most part, Sklerov's approach does not provide a satisfactory solution for the "response crisis". It works for the past cases of hacker attacks by Russia and China because both states acknowledged that the attacks were led from within their countries. Without such acknowledgement, however, the country of origin of a cyber attack cannot be determined. Any certain attribution of location has to be considered a systematic impossibility. Yet it is a necessary condition. The law of war cannot subsist without it, not even following Sklerov's interpretation.

---

[6] W.A. Owens, K.W. Dam & H.S. Lin (Eds.) *Technology, Policy, Law and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities*, Washington D.C. 2009

[7] Sklerov, p. 74.

[8] See Security Expert David Bianco, replying to Sklerov, available at: blog.vorant.com/2010/is-active-response-valid-approach-to.html