Ethics of Hacking Back

Six arguments from armed conflict to zombies

A policy paper on cybersecurity

Funded by: U.S. National Science Foundation

Prepared by: Patrick Lin, PhD
California Polytechnic State University
Ethics + Emerging Sciences Group
San Luis Obispo, California

Prepared on: 26 September 2016

Version: 1.0.0

Index

Abstract

It is widely believed that a cyberattack victim should not "hack back" against attackers. Among the chief worries are that hacking back is (probably) illegal and immoral; and if it targets foreign networks, then it may spark a cyberwar between states. However, these worries are largely taken for granted: they are asserted without much argument, without considering the possibility that hacking back could ever be justified. This policy paper offers both the case for and against hacking back—examining six core arguments—to more carefully consider the practice.

Acknowledgements

This policy paper has benefited from reviews by and conversations with Duncan Hollis, Heather Roff, Fritz Allhoff, Keith Abney, Rob Morgus, Peter Singer, and others. This research is supported by U.S. National Science Foundation grant #1318126. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the persons or organizations above.

01

Introduction

In cybersecurity, there's a certain sense of helplessness—you are mostly on your own. *You* are often the first and last line of defense for your information and communications technologies; there is no equivalent of state-protected borders, neighborhood police patrols, and other public protections in cyberspace.

For instance, if your computer were hit by "ransomware"—malware that locks up your system until you pay a fee to extortionists—law enforcement would likely be unable to help you.[1] The U.S. Federal Bureau of Investigation (FBI) offers this guidance: "To be honest, we often advise people to just pay the ransom," according to Joseph Bonavolonta, the Assistant Special Agent in Charge of the FBI's CYBER and Counterintelligence Program.[2]

Do not expect a digital cavalry to come to your rescue in time. As online life moves at digital speeds, law enforcement and state responses are often too slow to protect, prosecute, or deter cyberattackers. To be sure, some prosecutions are happening but inconsistently and slowly. The major cases that make headlines are conspicuously unresolved, even if authorities confidently say they know who did them.

Take, for example, the 2015 data breach at U.S. Office of Personnel Management: personnel records for more than 20 million federal workers were stolen, including sensitive background information for security clearances. Or think of any number of high-profile incidents. For the most part, there have been no arrests, no prosecution, no restitution—in essence, no satisfaction or justice for victims.

In that vacuum, it is no wonder that self-help by way of "hacking back" has been gaining attention.[3] Hacking back is a digital counterstrike against one's cyberattackers. Where law enforcement would warn us to not chase down a robber or retaliate against a criminal gang in the physical world, they naturally reject hacking back as a sound strategy in the cyber world.

But what exactly is the case against hacking back? While the question appears in the media, actual sustained arguments are hard to find. It is supposed to be obvious that civil society should reject the practice as illegal and unethical. This policy paper will explore both the general case for and against hacking back. This is important, since more response-options are needed to deal with growing threats.

Without laying out the arguments, critics could be ruling out the option too quickly.

I will focus on *general* arguments, because the specific context may make a difference in judging particular cases. For example, it matters whether a cyber counterstrike is proportionate, discriminate, and safeguarded against excessive collateral damage.[4] If it is not, then it may be immediately unethical, if not illegal.

This paper will also focus primarily on ethics. While the *legal* risks are large, the law is still unsettled, as there has not been a clear test-case for hacking back yet. When the law is unclear and needs to be clarified, it is useful to return to ethics—to go back to "first principles"—to help guide the law's evolution. This general ethics discussion, then, sets the stage for further conversations about law and policy, which are separate but related issues. If hacking back is generally unethical, that may make conversations about wisdom and legality moot. But if it is not clearly unethical, the wisdom and legality of the practice can be a productive study.

1.1 What is hacking back?

Hacking back sometimes goes by the euphemism of "active cyber defense."[5] The idea is to emphasize that this kind of hacking is not an unprovoked first strike but a counter-response to an attack, in case there is an ethical and legal difference between first and second strikes. But hacking back, even if defensive, is offensive in nature: it is a directed attack back at an aggressor, not just a protective block. If defense against an attack is holding up a shield, then "active" defense is wielding that shield as a weapon to harm, not only to absorb an attack. So, the euphemism is a bit of a misnomer and blurs the lines between offensive and defensive measures, in case there is an ethical and legal difference between those as well.

Hacking back can take many forms, nearly as diverse as hacking in the first place. An organization, for example, can collect information or trace the theft back to a particular system, that is, attribute the attack to a perpetrator. It can even take a next step of breaking in to delete or retrieve the stolen data. It can also activate the attacker's webcam and send back photos for evidence. Alternatively, the hack-back can be more serious, such as embedding your sensitive data with malicious code that locks down a cyber-thief's computer, as ransomware does. It can also corrupt the system files of a computer or network to make it permanently inoperable.

Because there are many ways you could hack back, they involve different levels of harm, from privacy intrusions to data breaches to physical damage. It also may matter *who* does the hacking back: a private individual who hacks back without the approval of law enforcement is more troubling than a state that hacks back on behalf of a victim. Therefore, some forms of counterattacking may be more problematic than others.

In this report, by "cyberattacks", I mean those that threaten confidentiality, integrity, or availability of a system—serious attacks that

would qualify as computer crimes and acts of hacking. In contrast, verbal attacks or defamation by electronic means are not cyberattacks in this paper. Cyberattacks also do not have to be harmful *per se*, but they at least commit wrongs. For instance, an unauthorized peek at your online diary might not harm you, but you were still wronged when your privacy was violated.

For this policy paper, I will have the hard cases in mind, such as hack-backs by private actors that do physical damage without much provocation; for instance, if the initial cyberattack had only shut down access to a non-critical website for even just a few minutes. If those cases are not generally unethical, then neither are the less troubling cases.

1.2 What is the controversy?

Unclear legal status is the root of hacking back's controversy. It is "probably illegal," as news reporting usually notes.[6] Looking at the U.S. as an example, the Department of Justice calls it "likely illegal" in its latest advisory for victims of cyberattacks.[7] The FBI "cautions" victims against hacking back but stops short of forbidding it.[8] At the highest level of government, White House officials call hacking back "a terrible idea."[9]

The same laws that make it illegal to hack in the first place—for instance, to access someone else's system without authorization—*presumably* make it illegal to hack back. In the U.S., the Computer Fraud and Abuse Act and Wiretap Act are among the key pieces in this patchwork of law. Foreign laws may be violated, too, such as the Computer Misuse Act and Data Protection Act in the U.K.; and the Budapest Convention on Cybercrime attempts to harmonize these and other such laws internationally.

However, these laws were not written with hacking back in mind: they do not consider hacking back, as distinct from unprovoked or standalone hacking more generally, and there is not yet a clear test-case to settle the question of whether or not the practice is legal. One reason for the lack of a test-case is a lack of prosecution of those who hack back, in any of the forms it may take. If initial cyberattacks are difficult to attribute or prosecute, then so are counterattacks.

Very few, if any, organizations admit to conducting such legally questionable actions, though some anonymously say that hacking back happens.[10] States may be reluctant to prosecute anyway, given a delicate relationship with industry, which is stressed under state demands for greater information-sharing.[11] As former U.S. Department of Justice attorney Bob Cattanach surmised, "The government's relationships with the private sector are so fragile that the Justice Department would probably exercise prosecutorial discretion and not bring a case to avoid damaging those ties."[12]

This is to say that, except for reckless cases of hacking back that hit innocent targets, it would be odd—and politically brave—to prosecute an individual or organization engaged in the practice, without also prosecuting the offender

who hacked first.[13] The cooperation of the initial offender is needed in order to have an actual victim to build a case against the counterattacker, and this is unlikely.[14] Authorities may also be turning a blind eye to certain kinds of hacking, as they need all the help they can get against common adversaries.[15]

At the same time, calls are increasing to consider hacking back as a response-option, even from the state itself. Without prosecutions or other public progress against cyberattackers, there is a temptation to strike back at the perpetrator, to achieve some measure of justice and deterrence.[16] Indeed, in its 2015 report back to Congress, the United States-China Economic and Security Review Commission recommended that:

> Congress assess the coverage of U.S. law to determine whether U.S.-based companies that have been hacked should be allowed to engage in counterintrusions [i.e., hacking back] for the purpose of recovering, erasing, or altering stolen data in offending computer networks. In addition, Congress should study the feasibility of a foreign intelligence cyber court to hear evidence from U.S. victims of cyberattacks and decide whether the U.S. government might undertake counterintrusions on a victim's behalf.[17]

Stewart Baker, former general counsel of the U.S. National Security Agency (NSA) and former assistant secretary for policy at the U.S. Department of Homeland Security (DHS), has been one of the most prominent advocates for hacking back. In 2013 Congressional hearings, he testified:

> We face a crisis. Cybersecurity is bad and getting worse. Civilian lives, our economic future, and our ability to win the next war, depend on solving our security problems. We need to find ways to turn the tables on hackers by putting the pressure on them and the entities that sponsor and enable them. To do this, we need to shift to a more active defense posture—one that relies on attribution and retribution. In my view, this shift would be best achieved if we find ways to allow victims to use their own resources, under government oversight, to identify the people who are stealing their secrets and the institutions that are benefiting from the theft.[18]

And a few months later, he noted:

> We will never defend our way out of the cybersecurity crisis. I know of no other crime where the risk of apprehension is so low, and where we simply try to build more and thicker defenses to protect ourselves…Sometimes the best defense is really a good offense; we need to put more emphasis on breaking into hacker networks...if we want a solution that will scale, we have to let the victims participate in, and pay for, the investigation. Too many government officials have viewed private countermeasures as a kind of

> vigilante lynch mob justice. That just shows a lack of imagination.[19]

This policy paper, then, responds to these and other calls to imaginatively consider hacking back—entertaining the case both for and against it, rather than merely presuming one way or another. The law governing this issue is still murky, as top experts continue to disagree on the subject.[20] Thus, our discussion here will abstract away from that legal quagmire and focus on the *moral* foundation underlying the debate.

02

Six arguments

The dominant expert opinion is that hacking back is wrong and should not be permitted. For instance, both *Bloomberg Business*[21] and *The Christian Science Monitor*[22] found that only fewer than 20% of its survey respondents supported hacking back. An expert panel at the 2016 RSA Conference, one of the world's largest cybersecurity meetings, agreed that hacking back was a bad idea.[23]

Given this dominant view, the opposing case has been underexplored. While I will not take a side in this debate—only offer analysis as neutrally as possible—much of the discussion will be to fill this gap by better developing the case for hacking back, as the case against it is already more intuitive.

In approaching this study, we first recognize that new technologies bring new dilemmas. For instance, with the development of nonlethal weapons such as the Active Denial System, otherwise known as the "pain ray", security personnel can now have an option between shouting and shooting.[24] Though it seems much better to temporarily cause pain than to kill, using the weapon for crowd control seems to violate a bedrock rule of war to never target noncombatants; and so the nonlethal weapon has yet to be deployed.

Likewise, cyberspace presents novel issues. To grapple with them, we often look toward the familiar—we look for analogies—for help. This is in large part how legal reasoning works in novel cases.[25, 26] But the challenge is that different analogies can frame the conversation in conflicting ways.[27] We see these competing narratives in ongoing debates about cyber norms, where cyber has been compared to outer space, international waters, Antarctica, the Wild West, and more.[28] The specific analogy or context matters, because different laws and norms may apply.

For instance, in a law-enforcement frame, tear gas may be used against attackers; but in a military frame, tear gas is prohibited by the Chemical Weapons Convention. As it applies here, looking at cyberoperations through different lenses can implicate different legal regimes, which carry different permissions and obligations.

Hacking back, in particular, has been framed in myriad ways and has been the subject of intense political, economic, legal, and media debates.[29, 30, 31, 32] In this paper, rather than focus on any given analogy, I will tease out the core arguments at play, both for and against hacking back. They are still intertwined to

some degree, as some are made in response to others, but each can stand on its own, and some are perhaps stronger than others.

The six arguments below are based on the rule of law, self-defense, attribution, escalation, public health, and practical effects.

2.1 Argument from the rule of law

At the conceptual level of the debate, it may be asserted that only the state has a legal monopoly on violence. Therefore, those non-state actors who hack back are exercising power that they do not legitimately have, and this erodes the rule of law.

The idea is that the state or government is based on a "social contract" among citizens to renounce their natural liberty to use force against one another; they transfer that power to the state itself.[33] Otherwise, without a higher power to appeal to, people would live in a perpetual condition of conflict. You could be attacked at any time, by any person who s/he believes has been aggrieved by you. Political philosopher Thomas Hobbes wrote:

> In such condition there is no place for industry, because the fruit thereof is uncertain, and consequently no culture of the earth, no navigation nor use of the commodities that may be imported by sea, no commodious building, no instruments of moving and removing such things as require much force, no knowledge of the face of the earth; no account of time, no arts, no letters, no society, and, which is worst of all, continual fear and danger of violent death, and the life of man solitary, poor, nasty, brutish, and short.[34]

Fast-forward to the modern world, to unilaterally decide to hack back would seem to break this social contract—to use force that we had promised to leave up to the state—and this undermines the basis for civil society. In other words, if it is indeed illegal, then hacking back openly flaunts the law, which may encourage the disregard of other laws. And civil society cannot function or promote justice without general respect for and compliance with the law.

On the other hand, if there is a social contract to swap our natural executive powers for collective security—a reasonable arrangement—it seems premised on the ability of the state to live up to its purpose of protecting us. If the state fails in this duty with respect to a particular threat, the entire social contract is not necessarily voided, but the state's monopoly on violence could be apportioned back to citizens to defend ourselves.

A common response at this point is that hacking back goes beyond self-defense and smacks more of vigilantism. If hacking back is defensible at all, then self-defense must be distinguished from vigilantism. That discussion is related to the argument from the rule of law, since vigilantism seeks to operate outside of it. (But I will address it in the next argument based on self-defense.)

Undoubtedly the rule of law is vital for civil society, but this argument from the rule of law may overstate the risk. It arguably commits a slippery-slope fallacy, in that not every violation of the law erodes the fabric of society. For instance, driving faster than the speed limit or illegally downloading music does not push the state toward anarchy. Hacking back, of course, can be a more serious violation but perhaps still not the spectacular violation that it's imagined to be.

Cyberspace is often considered as its own domain—as land, air, and sea are considered distinct domains—and the law's reach may exist to different degrees across domains. This may affect the individual's ability to exercise executive powers, such as to punish or defend. For instance, reprisals by individuals and companies against foreign targets were once permitted on difficult-to-protect seas. These private reprisals were authorized by "letters of marque" from a bygone era. With these letters, the state had empowered its citizens to repel pirates and foreign enemies on the open seas.[35] This practice, with obvious inherent risks, was ended by the Treaty of Paris in 1856, and the same rationale would seem to weigh against hacking back today.[36, 37]

Cyberspace, however, is arguably not ruled by law as international waters are today, given the apparent lack of law enforcement in the important cases at least. If this is right, then hacking back might not be a threat to the rule of law at all, because the law does not reach that far in cyberspace. Thus, we can look at cyberspace as more a "state of nature"—a state *prior* to civil society and government—that we eventually wish to tame.[38] While cyberspace springs forth from civil society, it transcends the physical borders of that cradle, resembling more a primitive state with fewer rules. In this emerging cyber state, if civil society is unable to uphold its end of the social contract, executive powers may revert back to the individual.

For some political theorists, a state of nature is a state of war. Hobbes declared, "Where there is no common power, there is no law; where no law, no injustice."[39] That is, anything goes in the state of nature. If cyberspace were a state of nature like this, then cyberattacks and counterattacks are neither ethical nor unethical. But life in Hobbes' state of nature is "solitary, poor, nasty, brutish, and short", driving us to escape that endless war by forming a social contract with others to establish rules.[40] This is something like a cyber-treaty proposal for nations.[41]

Other theorists are not as pessimistic. John Locke declared, "But though [the state of nature] be a state of liberty, yet it is not a state of licence."[42] That is, we are physically free, but not morally free, to do as we like. Reason dictates that we ought to be guided by, first, our duty to protect and sustain ourselves and then, when not in conflict with our own survival, our duty to protect and sustain others.

Cyberattacking, then, is wrong because it generally contravenes this second duty. Locke also allows that we may use lethal force to defend our property if a thief is using force, but not when the thief is fleeing and no longer a threat.[43] For Locke, the ethics of hacking

back may depend on whether the cyber-thief is still "in your house"—if that is a good analogy at all—and other particulars, to be discussed more in following sections.

Whichever theory we like about the state of nature, there is room for debate that the individual retains some natural liberty to exercise executive powers, such as to use force against active threats.

2.2 Argument from self-defense

The most important argument that supports hacking back is based on a natural right to defend one's own person and property.

Imagine that intruders break into your house in the dead of night. You do not know their intentions but have no reason to believe they will not harm or steal from you, especially after showing a disregard for basic criminal laws on breaking and entering. Would it ever be appropriate for you to attack the intruders?

According to most criminal laws, it would depend on the circumstances. If the intruders were fleeing or have fled, you probably do not have cause to hunt them down, as they are no longer a direct threat. If the intruder were a child who did not know any better, or maybe just someone who was desperate for food but was not aggressive, you do not have compelling reason to use force. If you could easily flee the scene and escape any danger, some would argue that you have a duty to retreat, saving deadly force as a last resort.

However, when your safety is in question, and even perhaps only your property, you seem to have a right to self-defense. In some U.S. states, this is codified as "stand your ground" laws and "castle" laws that limit or deny the existence of your duty to retreat.[44, 45]

This right, however, is less clear as an international matter. As the United Nations' Universal Declaration of Human Rights (UDHR), article 3, affirms: "Everyone has the right to life, liberty, and security of person."[46] And in UDHR, article 12: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."[47]

By themselves, those UDHR articles do not directly imply a right to self-defense. They only imply the state's duty to safeguard our right to life, security, and so forth—that is, your right to be *defended by* the state. This is relevant to the previous argument from the rule of law: even where your safety is in question and even inside your own home, the right response may be to summon law enforcement to your aid and not to take matters into your untrained hands, especially when it requires a judgment to use lethal force.

However, our intruder scenario is not a perfect analogy to a cyberattack, even a serious one. [48] One key difference between cyber intrusion and a home intrusion is that law enforcement is unlikely to rescue you in the event of a cyberattack. This inability is partly due to the difficulty of discovering and attributing cyberattacks, in order to know how to timely

respond. (See next argument based on attribution.)

Conceivably, in the absence of law enforcement, the individual retains the right to self-defense. Indeed, the state's duty to defend you appears to be derived from your prior right to self-defense, which you had traded for collective security.[49] As suggested in the preceding argument, when the state is unable or unwilling to meet that basic duty, the individual may become responsible once again for his or her own defense.

In ordinary scenarios, such as in a fistfight, you also seem to be legally and morally permitted to counterattack when the state cannot intervene quickly enough to protect you from harm; therefore, this is not an unusual principle. Even when police officers are one minute away, a lot of harm can happen to you in that minute. Thus, it is not unreasonable to defend yourself even where reliable law enforcement exists.

Yet cyberattacks are not fistfights or home invasions. It is often unclear who the attacker is, what level of harm they pose, if they're still inside your network, and so on. All this may determine how much force you are morally permitted to use, if any. Inasmuch as cyberattacks do not usually pose existential threats—though some have shuttered companies[50, 51]—the harm seems to be loss of confidentiality, integrity, or availability to systems and data, but not loss of life.

To assume a cyberincident is so serious that it warrants lashing back in that direction is to risk the use of disproportionate force, as well as targeting an innocent party; and these are generally regarded as unethical. Again, the specifics matter, and this policy paper examines mainly the *general* case for and against hacking back. At least some of the arguments here may permit the practice even in response to trivial attacks, which are the hard cases.

To be clear, there are sensible limits to defending one's own property, such as a home. Resisting a threatening intruder is one thing, but setting up automatic booby traps in one's home or on one's land is another. The automatic nature of such a response is objectionable because it bypasses any attempt at attribution; it could be that a trespasser is a lost child or another nonthreatening innocent.[52] A lethal booby trap, such as a spring-loaded shotgun, is worse and prohibited, as it puts more value on property than human life.[53]

Hacking back may be automatic or even autonomous, but it seems relevantly different from a booby-trap case. Though cyberattacks could theoretically lead to fatalities—such as hacking into power grids or medical devices—none has yet occurred; and a fatality would be a very rare exception in hacking back, not the rule nor its intent. Also, cyberintruders may be presumed to be unfriendly, as children and other innocent visitors are not penetrating cyberdefenses unwittingly, especially of an organization that takes security so seriously that it would contemplate hacking back.

A final objection to the argument from self-defense is that it seems to resemble vigilantism, which is generally regarded as unethical and illegal. Even if it does not erode

the rule of law, vigilantism risks escalating violence. (See the argument below based on escalation.)

However, we should be careful to distinguish vigilantism from self-defense. The latter is meant primarily to protect yourself from clear and present danger, especially when no help is immediately available. Vigilantism, in contrast, can be viewed as extrajudicial or gratuitous action, that is, taking action when judicial remedy, even if imperfect, is reliably available or using unnecessary force in self-defense.[54, 55]

While there are laws against cyberattacks, there's not a reliable judicial process for prosecution, deterrence, and restitution to which the victim can appeal. This means hacking back is not clearly a case of vigilantism. Whether a hack-back is a disproportionate use of force depends on the initial attack and countermeasure taken, but it's also not clear that hacking back is *always* an inappropriate use of force that should be categorically banned.

Reasonable self-defense, on the other hand, is permissible under nearly every ethical and legal theory, particularly when the state cannot intervene in time to prevent or defend against the initial attack. In these scenarios, fighting back—say, against a mugger or rapist—is a well-established exception to a general prohibition on harming others.

Further, reasonable self-defense does not appear to be illegal or wrong in the first place: "Justification does not make a criminal use of force lawful; if the use of force is justified, it cannot be criminal at all."[56] If no law is broken, then the rule of law cannot be threatened.

If hacking back might be permissible self-defense, more investigation is warranted on additional questions, such as whether self-defense can be claimed if the defense occurs much later, after the threat has disappeared. Is there an analogy to self-defense in a physical war, in which a delayed response is acceptable, e.g., in order to negotiate or for strategic reasons, because the threat is persistent? And what are the implications for a counterattack in cyberwar, long after an initial attack had occurred; by what measure do we judge whether a cyber threat is persistent?

At worst, even if self-defense is not justified in a cyberattack, we can still see it as an act of civil disobedience, possibly of the nonviolent kind, depending on the countermeasure taken. Victims do not want to resort to hacking back: it can be resource-intensive, legally risky, and potentially escalatory. But when there is no reasonable recourse, laws that foreclose the only remedy available—of self-defense—seem to be unhinged from reality. Breaking the law here can be a legitimate act of protest to change those laws.

2.3 Argument from attribution

Even if using force against an attacker is justified, whether in the physical world or cyber, there is a basic issue of attribution: you need to know who is attacking you to ensure you are not retaliating against an innocent

party, which would seem fundamentally unjust.

As recently as 2012, conventional wisdom suggested that attribution in cyberspace is notoriously difficult.[57,58] Even now, challenges remain, and the core problem with attribution in cyberspace revolves around two key components: time and certainty. Attribution comes in sequential increments that provide a growing level of certainty over time. Immediately after an intrusion is detected, for example, technical evidence can be compiled, which sometimes suggests a culprit.

Using this type of information, organizations in the private sector and government alike are often able to quickly attribute attacks with a very low degree of certainty. Follow-up investigations, which often unfold far more slowly, and the infusion of cyber threat intelligence—essentially profiles on different cyber threat-actors that have been identified previously—help identify culprits with far more certainty.[59]

But even with these tools, though, attribution is never fully certain, and the amount of time it takes to achieve highly certain attribution makes it unlikely that potential hackers-back will have proper attribution when hacking back; and without proper attribution, a counterattack might not be aimed at the guilty. Given that it is easy to route a cyberattack through innocent servers or spoof an IP address to hide the true origins of the attack, a counterattack with little or no oversight may hit only an innocent third-party—the messenger but not those responsible for the initial attack.

Botnets are an example of this problem. Infected systems are co-opted by the attacker who's able to control this horde of "zombie" computers for criminal purposes, such as to launch a distributed denial of service (DDoS) attack on a target site. The owners of these computers are usually unaware that their systems are compromised and being manipulated by unauthorized others. So, if the victim of a DDoS operation were to counterattack those hijacked computers, that would seem to damage the property of innocent users, which is usually to be avoided on legal and ethical grounds.

The moral argument against hacking back on the grounds that attribution is difficult relies on the assumption that clear attribution is a necessary pre-condition. Let's look critically at this assumption before accepting it: Is it really an ethical requirement to have clear attribution—to know that your target is the guilty party and the one truly responsible for your attack—before striking back? For instance, in a DDoS attack, is it always wrong to target the computers that were hijacked without the knowledge or consent of their owners?

Again, it is helpful to look at analogies to work through these novel cases. One analogy we can make is to an innocent attacker in the physical world, say, a random person who was coerced into committing armed robbery or a suicide-bombing. Can it ever be ethically permissible to attack this person who is morally innocent?

The answer seems to be yes. Even the police are not expected to ascertain this person's identity and motives before using lethal force against him or her, to prevent a worse outcome. We do not need to establish *mens rea* (Latin for "guilty mind") before we can act against a threat, or else it would always be too late. All that we need to know, at that moment, is that the person is a threat to others, culpability aside. So, it seems permissible or reasonable to use force against innocent people, at least in principle and under certain conditions.

To be fair, this analogy also is not quite right. In a cyberattack, it is not clear that the attacker remains at the scene to be a clear and present danger to anyone, unlike a coerced robber who is pointing a gun at others or an unwilling suicide-bomber. That is, there may not be the urgent need to counterattack, as there might be in analogous physical scenario.

Pursuing the cyberattacker, then, would seem to be more like pursuing and shooting at a robber who is fleeing. That is less a matter of self-defense and more a matter of justice—more appropriate for law enforcement to handle than private citizens.

Or is it? Remember, in cyber, there is an apparent lack of law enforcement, insofar as the most significant cyberattacks are unresolved, and prosecution can be slow and uneven. The state does not intervene much; and, no matter the reason for its inaction, this arguably returns power from the state back to the individual. Also, it is possible that sometimes the victim *is* in a better position to judge attribution and to respond, while the cyberattacker is still online or otherwise at the scene of the crime. Unclear or slow attribution, in the victim's eyes, may be exactly what motivates the practice. If a lack of confident attribution is what hinders arrests and prosecutions, then it seems unreasonable to ask the victim to do nothing indefinitely, while little progress is made on the specific case.

Even if attribution is considered by some to be a "solved" problem, it often fails to lead toward prosecution, as if there were no attribution at all. This disconnect not only erodes public confidence that the rule of law exists in cyberspace, but also suggests that attribution is a red herring or irrelevant issue in this debate, if attribution does not make a practical difference. None of this, however, sanctions reckless counterattacks when no attempt of attribution or other due diligence has been conducted. The claim here is that *definitive*—that is to say, fully certain—attribution may be too strict a requirement, especially when a rapid reaction is needed.

2.4 Argument from escalation

This is one of the most serious practical criticisms of hacking back, so I will dedicate more time to developing and examining this particular argument.

As with attribution, the counterattacking individual or company is usually in a worse position than state authorities to judge the escalatory ladder—how the adversary might respond—and contain any escalation. This is a

particular concern when cyberattacks come from abroad. We may imagine them to be the opening volleys of a cyberwar, which could escalate into a physical or kinetic war.[60]

Knowing that a cyberattack originated from a foreign territory is not by itself a smoking gun that a foreign government was behind it.[61] It might be state-sponsored hackers, but it also could be patriotic hackers, hacktivists, or ordinary criminals operating in that territory. Or the cyberattack might not have started from that territory at all; again, its source could be spoofed to frame an innocent country, precisely to create a conflict for it.

Regardless of attribution, hacking back against a foreign target may be misinterpreted by the receiving nation as a military response from our state, to serious political and economic backlash. Even if not perceived as a military action from the state, a poorly timed hack-back could derail delicate relationships and negotiations with a competitor state. Again, these are matters that seem better left to the state, not to private cowboys.

However, even if hacking back were conducted by the state to defend private victims, this could be problematic. Assuming the most challenging case that a cyberattack counts as a "use of force" or "armed attack", if hacking back were a salvo in cyberwar, does it violate international laws of armed conflict (LOAC)?

As declared by the United Nations Charter, article 2(4): "All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations."[62] Nonetheless, it is also within the natural rights of the attacked nation to defend itself, possibly allowing for hack-backs. As the U.N. Charter, article 51 declares: "Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security."[63] So, inside a cyberwar, hacking back by the state could be permitted. But what about before a cyberwar has started: would hacking back exacerbate the conflict and trigger that war? If so, this is a worst-case scenario that we would be right to guard against.

Yet, international cyberattacks have been ongoing for more than two decades without too much escalation or actual war. It makes sense that cyberattacks are tolerated and not too incendiary if we view cyberspace as a *borderland*: it is an ephemeral, unfamiliar domain that slips between a purely informational world and the physical world.[64]

In this framing, it is unclear whether a cyberconflict threatens territorial integrity, because said integrity requires borders to be clearly defined or asserted, and the borders of cyberspace are hard to locate in the first place. While cyberspace has an underlying physical reality in the servers, routers, transoceanic cables, and so on, and these physical items may reside within sovereign territories, cyberspace itself is difficult to place within those territories.

In philosophy, this is the notion of an emergent property: a whole that is greater than the sum of its parts. For instance, a human brain is composed of neurons, synapses, axons, hormones, and so on, and these parts can be located in time and physical space. However, the mind or consciousness that emerges from the brain is transcendent; an idea, emotion, or memory is not "physical", as we commonly understand it, even if it has physical origins.

And so it is with cyberspace.[65] Where an offensive idea is not an attack on a person's brain, likewise we should not assume that a conflict in cyberspace is an attack on sovereign territories, though it can be if physical assets are damaged and possibly under other conditions. That assumption would be too quick, committing something like a fallacy of composition or a genetic fallacy: improperly ascribing a feature about a part to its whole or to its origins.

Exiting our philosophical detour, if we are to accept the construction of cyberspace as contested frontier, then the following legal case is relevant. In the International Court of Justice (ICJ) case of *Nicaragua vs. United States of America* in 1984, the court's judgment distinguished an armed attack from a "mere frontier incident"; the latter is a less serious attack or use of force that does *not* trigger U.N. Charter's article 51 to justify a counterattack.[66] However, that does not mean the victim cannot counterattack at all, only that it cannot invoke its right to self-defense.

How a "mere frontier incident" is different than the more provocative "armed attack" and "use of force" is crucial to understand, since many international disputes are marked by such low-intensity violence. But it remains an underexplored distinction, surprising given its implications and controversy.[67] In the ICJ judgment itself, only a passing mention was made to the distinction, with no further explanation.

Legal defensibility aside, the distinction seems relevant to ethics, at least. Consider this situation: If you were jostling through a crowd, you may get bumped by other people, both intentionally and not. Either way, though, that unwanted physical contact does not usually rise to the level of assault. It may cause you some harm, but it is not so serious an offense that you would be justified in punching back in the name of self-defense, as opposed to pushing back as a natural reaction. Likewise, your bumping into others does not make you an assailant, and it is usually not considered criminal to bump back.[68] This is something like a frontier incident: it takes place in an environment to which no one has a firm claim, and the conflict is limited in its scale and effects. This unsettled environment affects your claim to bodily autonomy or state sovereignty, enabling certain transgressions to be taken less seriously.[69] By nature, frontiers are fraught with tests and misunderstandings.

While the states involved in a frontier incident cannot legally claim national self-defense and formally authorize war, it is understandable that the state agents—such as the explorers or military scouts at the frontier—would want to defend their own lives and deter future attacks. This might be achieved by returning fire at anything shooting at them, without

escalating into war. So, personal self-defense could justify a counterattack, even if a state's sovereignty is not at stake.

Returning to hacking back, even if the initial cyberattack were foreign-based and raised the possibility of cyberwar or worse, a counterattack seems to be reasonably treated as a "frontier incident" and not necessarily escalatory. It can usually be seen as a use of force *short of* war.[70] The rules and borders of the cyber frontier are unclear, as are its governing authorities.[71] This cyber frontier lines up with what military observers call a "gray zone" of conflict, an emerging space between war and peace.[72] These gray zones "involve some aggression or use of force, but in many ways their defining characteristic is ambiguity—about the ultimate objectives, the participants, whether international treaties and norms have been violated, and the role that military forces should play in response."[73]

Thus, it is not a radical idea to treat some attacks, either kinetic or cyber, as conflicts that are short of war but certainly disrupting the peace. We are still working out the rules. In cyber particularly, most attacks can then be considered to be frontier incidents, insofar as their ultimate objectives, participants, cybernorms, and so on are unclear. One exception would be when a cyberattack is known to be an act of war, for instance, if the aggressor declared it as such or was otherwise engaged in a kinetic war with the victim, in which cyberspace is one of several channels used to strike at the adversary.

A possible objection at this point again invokes attribution, as a key thread across various arguments. The objection is that the parties involved in a cyberattack might not be state-sponsored. Rather, they will likely be private entities, as most cyberattack victims are today. Thus, warfare often will not be the right frame in thinking about cyberattacks, and in cases where it is not, the subsequent analysis may be led astray.

Hacking back, however, is a case where this difference does not make a difference. In fact, the objection makes the argument for hacking back even easier. The state would be even *less likely* to be implicated and dragged into war, if there is less reason to believe that the involved parties were state-sponsored. Instead, perhaps it becomes a business-ethics problem between private entities from different countries. Even in that frame, cyber could still be viewed as a market frontier—sometimes chaotic and unfriendly. Economic conflict and crimes, of course, can and have escalated into full-blown war. The East India Company and piracy in the Caribbean from the 17th to 19th centuries are examples of private actors that have helped to spark wars. So, the risk of escalation into war is not zero, as a historical matter.

There is reason to think that economic cyber conflict is different, however. Again, there is no example in cyber's short but aggressive history that escalation is likely. The lack of precedent is perhaps owed to the difference in stakes, at least, as death is usually not on the line in cyber conflict. Also, unlike physical conflicts across borders or in contested territory, the borders of cyberspace are quite unclear. Cyberattacks are invisible, less visceral, and therefore less provocative, even though the economic costs may be quite high.

Former NSA and USCYBERCOM chief Keith Alexander called the costs "the greatest transfer of wealth in history."[74] Still, that harm has not provoked a kinetic battle or even escalation of cyberattacks. This, however, is no guarantee that serious escalation or political retaliation will never happen.

Even if a conflict rises to the level of war, attribution does not seem to be a firm requirement, at least for the individual soldier in times of imminent danger. It is enough to know that a sniper is shooting you from a certain position, without first identifying who the sniper is, which side he is on, his intentions, or anything else.[75] Notably, even if private entities were the only ones involved, the state might not be fully released from responsibility. Cyber norms have been proposed to hold the state liable if it fails to stop cyberattacks originating from its territory or otherwise fail to fully cooperate in their investigation.[76, 77] But staying out of such frontier incidents at least creates distance between the state and the parties involved, to lessen the risk of cyber or kinetic war.

But in the frontier event, we can grant that the initial attacker did something wrong. Outside of boxing and other sports, attacking first is usually wrong. Rather, the relevant question is whether the victim may return fire or push back. In both cases where the instigator is and is not blameworthy, it seems that the answer is yes, at least under certain conditions. This retaliation certainly holds special risks when it involves a foreign-based attacker, since LOAC may apply. But even exceptions or nuances exist in LOAC, as the ICJ ruling in *Nicaragua* shows.

Even without appealing to self-defense, it may be enough to observe that frontier incidents are an inherent risk to frontiers. Bad things happen here, and pushing back is one of those unfortunate but natural responses. It would be better if frontiers were more orderly and governed by law, but they fall in the legal gap between civil society and war, which are governed by different legal regimes.

Curiously, many people continue to conduct business and store data online, despite the relentless waves of cyberattacks. Coupled with the state's inaction to prosecute, this speaks to the risky, frontier-like environment of cyberspace. Therefore, operating in this environment is very much an assumed risk, just as building a house in a lawless frontier assumes the higher-than-average risk of being attacked.

Now, the frontier analogy is not perfect, like the others. More work can be done here to explore its strengths. For instance, frontiers typically imply a desired outcome that one group eventually wins control over some or all of it. But not all actors have plans for cyberspace domination; it is often understood to be a shared commons. Plausibly, cyber is more a loose community of people who have different interests and play in different areas, looking to get along, and less a battleground for nations. Does this difference matter? Maybe not, as we do not need to make that implication in the frontier analogy for it to be useful.

Thus, criticizing hacking back on the grounds that it may escalate a conflict is too broad an

objection. Again, *any* case of self-defense could be accused of the same provocation. This seems to be victim-blaming, similar to faulting a mugging victim for additional injuries sustained or created elsewhere as a result of fighting back. A mugging victim who fights back may be *causally* responsible for additional injuries arising from that action: if the victim had not fought back, then those injuries would not have occurred. But this is different from being *morally* responsible or blameworthy, if the victim bears no fault in initiating the series of events or does not use unreasonable force. Likewise, hacking back could very well be the reason why further retaliation and mayhem occurred, but those who hacked back are not necessarily to blame for that escalation.

2.5 Argument from public health

While the law-enforcement and armed-conflict frames employed above are the most obvious and natural to use, given the language of attacks and counterattacks, other analogies can be helpful here but also hold radical implications. For instance, some experts suggest framing cybersecurity as a public-health issue, and this has intuitive appeal.[78] Cyberattacks are a scourge or plague upon the public good of online life, and we already talk about immunizations against computer viruses and other malware. A botnet and other self-propagating attacks resemble a communicable disease, infecting computers without their hosts' knowledge and manipulating those "zombie" computers as part of a swarming plague.

In the fight against this cyber-zombie apocalypse, we do not need to do anything as theatrical as severing the heads of our attackers. Hacking back against a botnet can be as simple and nonaggressive as pushing security patches onto infected computers, just as patients with a deadly virus could be forcibly treated or quarantined to prevent a contagion's spread. To be sure, these are powers typically reserved for health authorities, but the responsibility to hold the line against an outbreak may revert to us, when there is no cyber equivalent of Centers for Disease Control and Prevention.

If security patches and other countermeasures damage innocent but hijacked computers, that would be unfortunate, as would be putting down a hapless zombie, especially one who was a friend or relative. But it is debatable whether real harm would have been done. Compromised machines should not be used anyway, and untreated machines (and zombie-bitten victims) are the reason why infections rage on.

A critic may reply that, in a pandemic, we should worry more about defense than offense; we ought to stay indoors and secure our own homes, instead of go outside to battle the infected.[79] This might be the safest course of action, but very few homes are self-sustaining and do not require contact with the outside world. Most organizations conduct business, and that means being connected to the outside world. Insisting that they should not interact with others is not a plausible option and again borders on victim-blaming.

While residents ought to take all reasonable precautions to secure their doors, no security is perfect. Sometimes pathogens can slip inside and ought to be dealt with quickly before they metastasize, especially if no one is coming to their rescue. Moreover, hacking back is not just for the benefit of the immediate victim, but it can be seen as a service to the public good. Just as we would want to stamp out every single smallpox virus, even those not in our vicinity or even if we were immune to the disease, targeting every single malevolent actor in cyberspace also better secures the community.

In medicine, if someone fails in this prevention and ends up on the receiving end of forced treatment, they are at least partly to blame for their fate, which can be as serious as unwanted quarantine.[80] At least where botnets are concerned, it may be that the co-opted computers are not truly "innocent" and therefore do not have moral immunity to counteraction. After all, malware is or should be a well-known risk for computer users by now.

Here, we should take notice of a couple things. First, this claim that computer users may bear some responsibility for their fate seems contrary to previous concerns about "victim-blaming." This suggests that the argument from public health, or at least the moral responsibility portion of it, may be mutually exclusive with other arguments that reject victim-blaming.

Or it could be that the concept of victim-blaming is ill-defined and needs more clarification. For instance, robbery victims who had left their homes or cars unlocked are usually thought to bear some responsibility; we can rightly blame them for enabling the crime. Maybe our world should not be so threatening as to impose such duties on us, but the reality is that it is. This clearly is a controversial and complex debate, beyond the scope of this paper.

Second, the public-health frame for cybersecurity may justify some actions that cannot be justified in other frames. As a public-health issue, by targeting any and all infected computers in a botnet, we are implying that it is permissible to target an entire class of undesired cyber actors, not just the immediate threats. But as a law-enforcement issue or frontier incident, self-defense permits us to strike back at only the immediate attacker, not against all criminals or soldiers of a competitor state. Thus, the public-health frame may hold broad implications to be determined.

Beyond the ethics of hacking back, our discussion here also suggests another analogy: computers—or the Internet itself—as dual-use products.[81] This was already implied in the law-enforcement and armed-conflict frames, and even in the zombie-apocalypse frame, where computers and networks are weapons. Misuse and irresponsible use of computers is the threat, with some pundits believing that the "Internet, whatever its many virtues, is also a weapon of mass destruction."[82]

The current lack of respect for the power of our computing devices is in large part what creates the debate on hacking back and other security issues. In fairness, the Internet was not designed for security when it was created

decades ago; it was made for only a small group of researchers who trusted one another.[83] That circle of trust is long gone, and now more vigilant and prepared users—and only very few of us are—are needed to prevent cyberattacks from landing in the first place, making moot the decision to hack back.

Therefore, to address cybersecurity at its human-factor roots—as we are often the weakest link—we may need to seriously consider special training and licensing for computer users, that is, requiring basic hygiene. Firearms and automobiles, likewise, have legitimate uses but also high potential for misuse, so they require proper training and licensing, too. The U.S. Federal Aviation Administration just required aerial drones to be registered, again recognizing both helpful and harmful uses.[84]

2.6 Argument from practical effects

Finally, there is the lingering issue of whether a counterattack would be practical or effective at all. If it is not, then it is unclear why hacking back should be permitted, especially given the risks.

If meant as a deterrent, hacking back would likely not deter malicious and ideological attackers. It might dissuade cybercriminals by imposing higher costs to their attacks, but it probably has little effect on attackers who are not motivated by profit and costs.

If meant as a remedy to the harm or wrong caused by the initial cyberattack, how exactly does one steal a secret back? Even if a victim were to break into his attacker's system and could locate his stolen data, deleting the file does not restore normalcy or security. There is no assurance that more copies of that data do not exist, and that data should be treated as permanently compromised. Furthermore, counterattacks do not fix holes when trust is lost in the integrity and security of one's system.

As mentioned previously, few organizations have the resources and technical capabilities to conduct an effective counterattack as a state could, and by preempting a state response to a cyberattack, hacking back may destroy evidence needed to make an arrest or prosecute the initial cyberattack.[85] These practical concerns are substantial but perhaps not insurmountable. For instance, hacking back is not advertised as a cure-all for cyberattacks. If it deters only rational or profit-motivated actors, that would be a good start and accounts for a majority of cyberattacks.

As far as remedy is concerned, if that were your only reason to hack back, then, yes, there may be no compelling reason to do it. Hacking back is not going to make you whole again. But deterrence and a basic sense of (retributive) justice—to make attackers suffer some negative consequence—are possible additional reasons. More holistically, if cybersecurity is a matter of public health, then killing a few zombies serves the larger public good.

It is also true that individuals and most companies are ill-equipped to mount a

counterattack by themselves. But as banks hire security guards, organizations could subscribe to "active defense" services made by cybersecurity firms. Or they could pool their resources and create security consortiums. However, there may be a question as to what extent hired cyber-guns are analogous to bank security guards and whether a right to self-defense is delegable.

Moreover, when outsourced to security firms, the business of hacking back raises another set of ethical questions not explored here, especially related to trust: Can you really trust others in giving them access to your organization's entire information system and data? Would there not be a conflict of interest with between treating a problem (ongoing revenue for your security firm) and curing it (which ends their engagement)? Would similar issues arise as with private military contractors?[86] And for many, allowing the state to counterattack on the victim's behalf may be worse than outsourcing to private security firms, insofar as they are less trusting of government with their sensitive data.

The worry about destroying evidence for prosecutors is a serious one. However, it also could amount to victim-blaming, even if a vague concept: we do not denounce resisting criminal acts on grounds that it might destroy evidence of the crime, or more broadly that the resistance would likely be ineffective. Of course, there may be prudential reasons to not fight back related to the victim's safety, but this is distinct from the odds of success. The concern over destroyed evidence seems ironic, given a low ratio of prosecutions to cyberattacks. Instead of prohibiting hacking back, a formal judicial process could give oversight to the practice, as any state assistance would be welcomed to bring light into this cyber underworld.

Currently, hacking back occurs in the twilight zone outside of the law, and the lack of prosecution for hacking back might be regarded as an endorsement of the practice; thus, the state could be implicated either way. Even if the law clearly prohibits it, hacking back may continue anyway out of desperation, as victims see little assistance from the state or prosecution of their tormentors.

Practicality and effectiveness are legitimate concerns about hacking back.[87] As with many other things, legalizing the practice can create opportunity for regulation to ensure that it is not abused, to strengthen the rule of law, and to help make it more effective. Due-process safeguards may include a streamlined process for *ex ante* judicial warrants (before a counterstrike) or *ex post* justification (after a counterstrike), if there is not enough time to seek a warrant.[88] Currently, there is no self-reporting of hacking back, because the practice is "likely illegal."

We cannot track what we do not measure. Without that data—a way for individuals and organizations to safely report countermeasures, without fear of being made into criminals—it is difficult to answer the question of whether hacking back has deterrent value, which is an empirical question. At any rate, openly doing nothing, as seems to be the case now, certainly offers no deterrence and likely encourages cyberattackers to continue preying on others.

If hacking back is currently ineffective, that may be more a problem with the legal environment, rather than with the practice itself. The rule of law is important, and the purpose of law is not only to prohibit but also to empower, as is the case in contract law and marriage law.

03

Conclusion

This is only the beginning of a systematic study of the issues, not the end of it. Strong intuitions exist against hacking back, but reasonable arguments also exist to support it. Nonetheless, as the saying goes, "in theory, there's no difference between theory and practice; but in practice, there is."[89] Even if we have some natural liberty to enforce justice—to hack back—there remains a rather large pragmatic risk of whether the state may prosecute this action anyway, theory aside. So, clarifying the law on hacking back is important to align theory with practice.

Like having to defend yourself in any conflict, whether offensively or defensively, hacking back is far from the ideal response. But it could be ethically permitted as a stop-gap measure, until cybersecurity and law enforcement are better able to identify and prosecute attackers. Testifying again before Congress, former NSA general counsel Stewart Baker emphasized the role of state regulation in addressing the risks of this stop-gap measure:

> I understand the concern expressed by some that we cannot turn cyberspace into a free-fire zone, with vigilantes wreaking vengeance at will. No one wants that. Government should set limits and provide oversight for a true public-private partnership, in which the private sector provides many of the resources and the public sector provides guidance and authorities. The best way to determine how much oversight is appropriate is to move cautiously but quickly to find alternatives to the current failed cybersecurity strategy.[90]

If hacking back is not as morally perilous as it has been presumed, the conclusion of this policy paper is *not* that we ought to immediately authorize the practice. The next step is to more fully explore legal and policy challenges in hacking back[91], as they may have been dismissed too quickly as well.

In this paper, I am not taking sides on the debate, though I do spend more time building a provocative case for hacking back; the opposing view is already much more dominant and better developed. At a time when we need more response-options to cyber threats, and when we are still grappling with the cyber domain conceptually, it may be premature to take reasonable options off the table.

~ ~ ~

04

Endnotes

[1] Winton, Richard. 2016. "$17,000 bitcoin ransom paid by hospital to hackers sparks outrage." *Los Angeles Times*. February 19. <http://www.latimes.com/local/lanow/la-me-ln-17000-bitcoin-ransom-hospital-outrage-20160219-story.html>

[2] Roberts, Paul. 2015. "FBI's advice on ransomware? Just pay the ransom." *The Security Ledger*. October 22. <https://securityledger.com/2015/10/fbis-advice-on-cryptolocker-just-pay-the-ransom/>

[3] Pennington, Matthew. 2015. "U.S. advised to examine 'hack back' options against China." *Associated Press*. November 17. <http://bigstory.ap.org/article/79f4659ffb6346d0ac6d01f83c9987cc/us-advised-examine-hack-back-options-against-china>

[4] Jensen, Eric. 2013. "Cyber attacks: proportionality and precaution in attacks." *International Law Studies*, volume 89, issue 1, pp. 198-217. <http://stockton.usnwc.edu/ils/vol89/iss1/15/>

[5] Lachow, Irving. 2013. "Active cyber defense: a framework for policymakers." *Center for a New American Security*. February 22. <http://www.cnas.org/publications/policy-briefs/active-cyber-defense-a-framework-for-policymakers>

[6] Kuchler, Hannah. 2015. "Cyber insecurity: hacking back." *Financial Times*. July 27. <http://www.ft.com/intl/cms/s/2/c75a0196-2ed6-11e5-8873-775ba7c2ea3d.html>

[7] U.S. Department of Justice. 2015. "Best practices for victim response and reporting of cyber incidents." Cybersecurity Unit, Computer Crime & Intellectual Property Section, Criminal Division. <http://www.justice.gov/sites/default/files/opa/speeches/attachments/2015/04/29/criminal_division_guidance_on_best_practices_for_victim_response_and_reporting_cyber_incidents.pdf>

[8] Riley, Michael and Jordan Robertson. 2014. "FBI probes if banks hacked back as firms mull offensives." *Bloomberg*. December 30. <http://www.bloomberg.com/news/articles/2014-12-30/fbi-probes-if-banks-hacked-back-as-firms-mull-offensives>

[9] Gaines, Amanda, John Williams, and Peter Singer. 2015. "A 'cyber party' with John McAfee and the White House Cybersecurity Czar." New America podcast, interview with Michael Daniel at 23:48. <https://www.newamerica.org/cybersecurity-initiative/a-cyber-party-with-john-mcafee-and-the-white-house-cybersecurity-czar/>

[10] Harris, Shane. 2014. "How corporations are adopting cyber defense and around legal barriers." *Slate.* November 12. <http://www.slate.com/articles/technology/future_tense/2014/11/how_corporations_are_adopting_cyber_defense_and_around_legal_barriers_the.single.html>

[11] Fung, Brian. 2015. "Apple and Dropbox say they're against a key cybersecurity bill, days before a crucial vote." October 20. <https://www.washingtonpost.com/news/the-switch/wp/2015/10/20/apple-says-its-against-a-key-cybersecurity-bill-days-before-a-crucial-vote/>

[12] Robertson, Jordan and Michael Riley. 2014. "Would the U.S. really crack down on companies that hack back?" *Bloomberg*. December 30. <http://www.bloomberg.com/news/2014-12-30/why-would-the-u-s-crack-down-on-companies-that-hack-back-.html>

[13] *Ibid*.

[14] Alexis, Alexei. 2013. "Debate brewing over whether companies should strike back at their cyber attackers." *Bureau of National Affairs*. April 9. <http://www.bna.com/debate-brewing-whether-n17179873246/>

[15] O'Connell, Justin. 2015. "Stanford scholar: U.S. unlikely to prosecute Anonymous for harassing ISIS." *Hacked*. November 24. <https://hacked.com/stanford-scholar-us-unlikely-prosecute-anonymous-harassing-isis/>

[16] Though both justice and deterrence are reasons for punishment, they're distinct concepts. Aristotle's notion of justice still holds today, that justice restores the moral and social equilibrium that existed prior to a crime, which can include penalties against the offender. This is related to John Rawls' idea of justice as fairness. But deterrence is aimed at preventing future crimes, imposing a cost to discourage behavior rather than remedy wrongs and harms. See, Feinberg, Joel. 2008. "The classic debate." *Philosophy of Law*, eds. Joel Feinberg and Jules Coleman, 8th edition.

[17] United States-China Economic and Security Review Commission. 2015. *2015 Report to Congress: Executive Summary and Recommendations*, page 27. <http://origin.www.uscc.gov/sites/default/files/annual_reports/2015%20Executive%20Summary%20and%20Recommendations.pdf>

[18] Baker, Stewart. 2013a. "The attribution revolution: raising the costs for hackers and their customers." Testimony before the U.S. Senate Judiciary Committee's Subcommittee on Crime and Terrorism. May 8. <http://www.judiciary.senate.gov/imo/media/doc/5-8-13BakerTestimony.pdf>

[19] Baker, Stewart. 2013b. "The Department of Homeland Security at 10 years: examining challenges and achievements and addressing emerging threats." Testimony before the U.S. Senate Committee on Homeland Security and Governmental Affairs. September 11. <http://www.hsgac.senate.gov/download/?id=AD53570C-3682-45DF-91B1-E1D60F003CF7>

[20] Baker, Stewart. 2012. "The hackback debate." Steptoe & Johnson. November 2. <http://www.steptoecyberblog.com/2012/11/02/the-hackback-debate/>

[21] Robertson, Jordan. 2015. "It's the government's job to respond to cyber attacks: Bloomberg poll." *Bloomberg*. January 21. <http://www.bloomberg.com/news/articles/2015-01-21/cyber-attack-retaliation-seen-as-government-s-job-in-global-poll>

[22] Sorcher, Sara. 2015. "Influencers: companies should not be allowed to hack back." *The Christian Science Monitor*. April 1. <http://www.csmonitor.com/World/Passcode/Passcode-Influencers/2015/0401/Influencers-Companies-should-not-be-allowed-to-hack-back>

[23] Armerding, Taylor. 2016. "Hacking back will only get you in more trouble." *CSO.* March 3. <http://www.csoonline.com/article/3040408/security/hacking-back-will-only-get-you-in-more-trouble.html>

[24] Lin, Patrick. 2013. "Pain rays and robot swarms: the radical new war games the DoD plays." *The Atlantic*. April 15. <http://www.theatlantic.com/technology/archive/2013/04/pain-rays-and-robot-swarms-the-radical-new-war-games-the-dod-plays/274965/>

[25] Lamond, Grant. 2006. "Precedent and analogy in legal reasoning." *Stanford Encyclopedia of Philosophy*. June 20. <http://plato.stanford.edu/entries/legal-reas-prec/>

[26] Sunstein, Cass. 1993. "On analogical reasoning." *Harvard Law Review*, volume 106, number 3, pp. 741-791.

[27] Sulek, David and Ned Moran. 2009. "What analogies can tell us about the future of cybersecurity." *The Virtual Battlefield: Perspectives on Cyber Warfare*. <https://ccdcoe.org/sites/default/files/multimedia/pdf/08_SULEK_What%20Cyber%20Analogies%20Can%20Tell%20Us.pdf>

[28] Goldman, Emily and John Arquilla. 2014. *Cyber Analogies*. <http://calhoun.nps.edu/bitstream/handle/10945/40037/NPS-DA-14-001.pdf>

[29] Paletta, Damian and Dion Nissenbaum. 2015. "Debate deepens over response to cyberattacks." *The Wall Street Journal*. Feb 8. <http://www.wsj.com/articles/debate-deepens-over-response-to-cyberattacks-1423434725>

[30] Arnold, Martin, Tom Brathwaite, and Hannah Kuchler. 2015. "Davos 2015: banks call for free rein to fight cyber crime." *Financial Times*. January 22. <http://www.ft.com/intl/cms/s/0/d94e855c-a209-11e4-bbb8-00144feab7de.html>

[31] Harrington, Sean. 2014. *"*Cyber security active defense: playing with fire or sound risk management*?" Richmond Journal of Law & Technology*, volume 20, issue 4. <http://jolt.richmond.edu/v20i4/article12.pdf>

[32] Lin, Patrick. 2015. "Should Washington allow companies to strike back against hackers?" *The Wall Street Journal*. May 10. <http://www.wsj.com/articles/should-washington-allow-companies-to-strike-back-against-hackers-1431022206>

[33] Hobbes, Thomas. 1651. *Leviathan*, at chapter 13. <http://www.gutenberg.org/files/3207/3207-h/3207-h.htm>

[34] Ibid, para. 9.

[35] Owens, William, Kenneth Dam, and Herbert Lin. 2009. *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities*. <http://www.nap.edu/download.php?record_id=12651>

[36] *Declaration Respecting Maritime Law*. 1856. Paris. April 16. <https://www.icrc.org/ihl/INTRO/105?OpenDocument>

[37] Roff, Heather. 2014. "Cyber letter of marque and reprisals: 'hacking back'." *Duck of Minerva*. October 15. <http://duckofminerva.com/2014/10/cyber-letters-of-marque-and-reprisal-hacking-back.html>

[38] Kaminiski, Ryan. 2010. "Escaping the cyber state of nature: cyber deterrence and international institutions." *Conference on Cyber Conflict Proceedings*, pp. 79-94. <https://ccdcoe.org/sites/default/files/multimedia/pdf/Kaminski%20-%20Escaping%20the%20Cyber%20State%20of%20Nature%20Cyber%20deterrence%20and%20International%20Institutions.pdf>

[39] Hobbes, Thomas. 1651. *Leviathan*, at chapter 13, paragraph 13. <http://www.gutenberg.org/files/3207/3207-h/3207-h.htm>

[40] *Ibid*, at paragraph 9.

[41] Calls for a treaty raise a set of questions not explored here, such as whether new laws are actually needed, as opposed to applying and enforcing existing law. See, Muller, Benjamin. 2014. "On the need for a treaty governing cyber conflict." *London School of Economic Ideas*. <http://www.lse.ac.uk/IDEAS/publications/reports/pdf/SU14_2_Cyberwarfare.pdf>

[42] Locke, John. 1689. *Second Treatise of Government*, at chapter 2, section 6. <http://www.gutenberg.org/files/7370/7370-h/7370-h.htm>

[43] *Ibid*, at chapter 3, section 18.

[44] Lin, Patrick. 2012. "'Stand your cyberground' law: a novel proposal for digital security." *The Atlantic*. April 30. <http://www.theatlantic.com/technology/archive/2012/04/stand-your-cyberground-law-a-novel-proposal-for-digital-security/256532/>

[45] These laws remain controversial, but our discussion here does not rely on them or assume they are ethical. They're offered only as an illustration of a formal right to self-defense.

[46] United Nations. 1948. "The Universal Declaration of Human Rights." December 10. <http://www.un.org/en/documents/udhr/>

[47] *Ibid.*

[48] The analogy is even weaker when we consider cross-border cyberattacks which have geopolitical implications. See the argument from escalation below for a discussion of self-defense in war or armed conflict at the state level.

[49] Cudd, Ann. 2012. "Contractarianism." *Stanford Encyclopedia of Philosophy*. August 2. <http://plato.stanford.edu/entries/contractarianism/>

[50] Brito, Helena. 2012. "Nortel breach raises questions about role industrial espionage played in company downfall." FireEye. March 27. <https://www.fireeye.com/blog/threat-research/2012/03/nortel-breach-raises-questions-role-industrial-espionage.html>

[51] Stahl, Lesley. 2016. "The great brain robbery." *CBS News*. January 17. <http://www.cbsnews.com/news/60-minutes-great-brain-robbery-china-cyber-espionage/>

[52] This is related to the emerging idea of "meaningful human control" over weapons. See, United Nations Institute for Disarmament Research. 2014. "The weaponization of increasingly autonomous technologies: considering how meaningful human control might move the discussion forward." *UNIDIR Resources*. <http://www.unidir.org/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf>

[53] These laws may vary in different jurisdictions but are generally consistent. For example, see *Katko v. Briney*. 1971. Iowa Supreme Court. 183 N.W.2d 657. <https://scholar.google.com/scholar_case?case=18983727091145856666>

[54] Johnston, Les. 1996. "What is vigilantism?" *British Journal of Criminology*, volume 36, issue 2, pp. 220-236. <http://bjc.oxfordjournals.org/content/36/2/220.abstract>

[55] How reliable the judicial process ought to be, before one may turn to extrajudicial measures, is the subject of another conversation.

[56] *The People of the State of New York v. Jerry McManus*. 1986. Court of Appeals of the State of New York. 67 N.Y.2d 541. <https://casetext.com/case/people-v-mcmanus-4>

[57] Lin, Patrick, Fritz Allhoff, and Neil Rowe. 2012. "War 2.0: cyberweapons and ethics." Communications of the ACM, volume 55, number 3, pp. 24-26. <http://cacm.acm.org/magazines/2012/3/146257/fulltext>

[58] For instance, Koh, Harold. 2012. "International law in cyberspace." USCYBERCOM Inter-Agency Legal Conference. September 18. <http://www.state.gov/s/l/releases/remarks/197924.htm>

[59] For more on attribution, see: Rid, Thomas and Ben Buchanan. 2014. "Attributing Cyber Attacks." Journal of Strategic Studies, volume 38. December 23. <https://sipa.columbia.edu/system/files/Cyber_Workshop_Attributing%20cyber%20attacks.pdf>

[60] It might be more commonly imagined that the cyberattacks are acts of espionage, that theft of intellectual property or secrets is the goal. In this case, escalation would not be much of an issue, as spying has not been, and is not, a *casus belli* or a cause for war.

[61] Schmitt, Michael, ed. 2012. *Tallinn Manual on the International Law Applicable to Cyber Warfare*. <https://ccdcoe.org/tallinn-manual.html>

[62] United Nations. 1945. "Charter of the United Nations." June 26. <http://www.un.org/en/charter-united-nations/>

[63] *Ibid.*

[64] Indeed, the existence of the "Dark Net"—as covered breathlessly in the media and research as a "wilderness", "Wild West", "underworld", and other ominous terms—is a testament to the shadowy nature of much of cyberspace. As the deep basement of the online world beneath the "surface Internet", i.e., websites that can be found by search engines, the expanse of the Dark Net is unknown and has been estimated to comprise 50% to 80% and more of all online content. See, Bartlett, Jamie. 2015. *The Dark Net: Inside the Digital Underworld*. Melville House. <https://www.amazon.com/Dark-Net-Inside-Digital-Underworld/dp/1612194893>. Langewiesche, William. 2016. "Welcome to the Dark Net, a wilderness where invisible world wars are fought and hackers roam free." *Vanity Fair*, October. <http://www.vanityfair.com/news/2016/09/welcome-to-the-dark-net>. Kushner, David. 2015. "The Darknet: is the government destroying 'the Wild West' of the Internet'?" *Rolling Stone*, October 22. <http://www.rollingstone.com/politics/news/the-battle-for-the-dark-net-20151022>

[65] Hunter, Dan. 2003. "Cyberspace as a place and the tragedy of the digital anticommons." *California Law Review*, volume 91, issue 2, pp. 439-519. <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1380&context=californialawreview>

[66] In this legal proceeding, the U.S. was charged with violating international law by supporting the revolution against Nicaragua at the time, but the details of this case are not relevant to our discussion. See, International Court of Justice. 1986. "Military and paramilitary activities in and against Nicaragua (Nicaragua v. United States of America) merits, judgment." *International Court of Justice Reports*, at paragraph 195. June 27. <http://www.icj-cij.org/docket/files/70/6503.pdf>

[67] Watts, Sean. 2011. "Low-intensity computer network attack and self-defense." *International Law Studies*, volume 87, pp. 59-87. <http://www.law.berkeley.edu/files/watts--low_intensity_computer_network_attack.pdf>

[68] To be clear, this is not meant to be an analogy to a cyberattack, but an example from everyday life or a metaphorical frontier incident, to show that the idea has intuitive appeal.

[69] There's no *mens rea* in the crowd analogy, but that difference seems to make a stronger case for firing back in the armed-conflict case where there is intent to harm you.

[70] Lin, Herbert. 2012. "Cyber conflict and international humanitarian law." *International Review of the Red Cross*, volume 94, number 886. <https://www.icrc.org/eng/resources/documents/article/review-2012/irrc-886-lin.htm>

[71] Melzer, Nils. 2011. "Cyberwarfare and international law." *UNIDIR Resources*. <http://www.unidir.org/files/publications/pdfs/cyberwarfare-and-international-law-382.pdf>

[72] Mazarr, Michael. 2015. *Mastering the Gray Zone: Understanding the Changing Era of Conflict*. United States Army War College Press. <http://www.strategicstudiesinstitute.army.mil/pubs/display.cfm?pubID=1303>

[73] Barno, David and Nora Bensahel. 2015. "Fighting and Winning in the 'Gray Zone'." *War on the Rocks.* May 19. <http://warontherocks.com/2015/05/fighting-and-winning-in-the-gray-zone/>

[74] Rogin, Josh. 2012. "NSA chief: cybercrime constitutes 'the greatest transfer of wealth in history'." *Foreign Policy*. July 9. <http://foreignpolicy.com/2012/07/09/nsa-chief-cybercrime-constitutes-the-greatest-transfer-of-wealth-in-history/>

[75] This is not meant necessarily to be an analogy with cyberattacks, but just to illustrate the claim that attribution might not be a firm requirement in war, which has implications for attribution in cyber.

[76] Tikk, Eneken. 2011. "Ten rules for cyber security." *Survival*, volume 53, issue 3, pp. 119-132. <http://www.tandfonline.com/doi/pdf/10.1080/00396338.2011.571016>

[77] United Nations. 2015. "Group of governmental experts on developments in the field of information and telecommunications in the context of international security." July 22. <http://www.un.org/ga/search/view_doc.asp?symbol=A/70/174>

[78] Singer, Peter and Allan Friedman. 2014. *Cybersecurity and Cyberwar*. <http://www.cybersecuritybook.com>

[79] Roff, Heather. 2015. "Should Washington allow companies to strike back against hackers?" *The Wall Street Journal*. May 10. <http://www.wsj.com/articles/should-washington-allow-companies-to-strike-back-against-hackers-1431022206>

[80] Latson, Jennifer. 2014. "Refusing quarantine: why Typhoid Mary did it." *Time*. November 11. <http://time.com/3563182/typhoid-mary/>

[81] Weaver, Nicholas. 2013. "Our government has weaponized the Internet. This is how they did it." *Wired*. November 13. <http://www.wired.com/2013/11/this-is-how-the-internet-backbone-has-been-turned-into-a-weapon/>

[82] Koppel, Ted. 2015. "Where is America's cyberdefense plan?" *Washington Post*. October 30. <https://www.washingtonpost.com/opinions/lets-talk-about-a-cyberdefense-plan/2015/10/30/efb19060-7cd7-11e5-b575-d8dcfedb4ea1_story.html>

[83] Timberg, Craig. 2015. "A flaw in the design." *Washington Post*. May 30. <http://www.washingtonpost.com/sf/business/2015/05/30/net-of-insecurity-part-1/>

[84] U.S. Federal Aviation Administration. 2015. "Press release – U.S. transportation secretary Anthony Foxx announces unmanned aircraft registration requirement." October 19. <https://www.faa.gov/news/press_releases/news_story.cfm?newsId=19594>

[85] Alexis, Alexei, *op. cit.*

[86] See, Singer, Peter. 2005. *Corporate Warriors: The Rise of the Privatized Military Industry.*

[87] Majuca, Roberto and Jay Kesan. 2010. "Hacking back: optimal use of self-defense in cyberspace." *Chicago-Kent Law Review*, volume 84, number 3, pp. 831-840. <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1363932>

[88] Lin, Patrick, Fritz Allhoff, and Keith Abney. 2014. "Is warfare the right frame for the cyber debate?" *The Ethics of Information Warfare,* eds. Luciano Floridi and Mariarosaria Taddeo. <http://www.springer.com/us/book/9783319041346>

[89] Savitch, Walter. 1984. *Pascal: An Introduction to the Art and Science of Programming*.

[90] Baker, Stewart 2013b, *op. cit.*

[91] For recent work on this, see, Rabkin, Jeremy and Ariel Rabkin. "Hacking back without cracking up." A Hoover Institution essay. <http://www.hoover.org/research/hacking-back-without-cracking>

About the author

Patrick Lin, Ph.D., is the director of the Ethics + Emerging Sciences Group at California Polytechnic State University, San Luis Obispo, where he is an associate philosophy professor. He is also affiliated with Stanford Law School's Center for Internet and Society, University of Notre Dame's Emerging Technologies of National Security and Intelligence Initiative, Australia's Centre for Applied Philosophy and Public Ethics, and the World Economic Forum's Global Future Councils. Previously, he held academic appointments at Stanford's School of Engineering, U.S. Naval Academy, and Dartmouth College.

Dr. Lin is well published in the ethics of emerging technologies—including robotics, cybersecurity, AI, human enhancements, nanotechnology, and more—especially their national security implications. He is the lead editor of *Robot Ethics* (MIT Press, 2012), among other books and articles. He has provided briefings, testimony, and counsel to the U.S. Dept. of Defense, CIA, United Nations, UNIDIR, National Research Council, Google, Apple, and many other organizations. He earned his B.A. from UC Berkeley and Ph.D. from UC Santa Barbara.

Contact: palin@calpoly.edu
Website: http://ethics.calpoly.edu

Ethics + Emerging Sciences Group