DOI:10.1145/2093548.2093558

Patrick Lin, Fritz Allhoff, and Neil C. Rowe

Computing Ethics
War 2.0: Cyberweapons and Ethics

Considering the basic ethical questions that must be resolved in the new realm of cyberwarfare.

By opening new channels for communication and services in a society, cyberspace also offers new opportunities for warfare. Indeed, it is more attractive than conventional military actions that require the expense and risk of transporting equipment and deploying troops in enemy territory, not to mention the political risk. Cyberweapons could be used to attack anonymously at a distance while still causing much mayhem, on targets ranging from banks to media to military organizations.

Today, many nations have the capability to strike in cyberspace—but should they? International humanitarian laws, or the "laws of war," have not been written with cyberspace in mind. So we face a large policy gap, which organizations internationally have tried to address in recent years, such as the U.S. National Research Council.[8]

But there is also a gap in developing the ethics behind policies. Ethics is an important branch of philosophy with a tradition of more than 2,000 years, and warfare has long been an important subject for ethics. Cyberwarfare challenges many assumptions of the traditional analysis of the ethics of warfare, so it is useful to examine cyberwarfare from an ethical as well as a policy perspective. In this column, we describe some new and previously identified issues related to ethics that need attention.

Aggression

By the laws of war, there is historically only one "just cause" for war: a defense to aggression. But since aggression is usually understood to mean that human lives are directly in jeopardy, it becomes difficult to justify military response to a cyberattack that does not cause kinetic or physical harm as in a conventional or Clausewitzian sense. Cyberattacks can be distinguished from cyberespionage by their deliberate damage; some clever cyberattacks can be subtle and difficult to distinguish from routine breakdowns and malfunctions, but usually the effect is obvious because it is intended to have political impact.

Does it count as aggression when malicious software has been installed on an adversary's computer systems that we believe will be imminently triggered? Or maybe the act of installing malicious software is an attack itself, much like installing a land mine? What about unsuccessful attempts to install malicious software? Do these activities count as war-triggering aggression—or mere crimes, which do not fall under the laws of war? Traditional military ethics would answer all these questions negatively, but they feature in debates over the legitimacy of preemptive and preventative war.[4]

Another important question to consider: Insofar as most cyberattacks do not directly target lives, are they as serious? The organized vandalism of cyberattacks could be serious if it prevents a society from meeting basic human needs like providing food. A lesser but still serious case was the denial-of-service cyberattacks on media-infrastructure Web sites in the country of Georgia in 2008, which prevented the government from communicating with its citizens.

Discrimination

The laws of war mandate that noncombatants be avoided in attacks, since they do not pose a military threat.[6] Most theorists accept a double effect in which some noncombatants could be unintentionally harmed as "collateral damage" in pursuing important military objectives,[2] though some have more stringent requirements.[13] Some challenge whether noncombatant immunity is really a preeminent value,[1] but the issue undoubtedly has taken center stage in just-war theory and therefore the laws of war.

It is unclear how discriminatory cyberwarfare can be: If victims use fixed Internet addresses for their key infrastructure systems, and these could be found by an adversary, then they could be targeted precisely—but victims are unlikely to be so cooperative. Therefore, effective cyberattacks need to search for targets and spread the attack; yet, as with viruses, this risks involving noncombatants. The Stuxnet worm in 2010 was intended to target Iranian nuclear processing facilities, but spread far beyond its intended targets.[10] Although its damage was highly constrained, its quick broad infection through vulnerabilities in the Microsoft Windows operating system was noticed and required upgrades to antivirus software worldwide, incurring a cost to everyone. The worm also provided excellent ideas for new exploits that are already being used, another cost to everyone. Arguably, then, Stuxnet did incur some collateral damage.

Cyberattackers could presumably appeal to the doctrine of double effect, arguing that effects on noncombatants would be foreseen but unintended. This may not be plausible, given how precise computers can be when we want them to be. Alternatively, cyberattackers could argue that their attacks were not directly against noncombatants but against infrastructure. However, attacking a human body's immune system as the AIDS virus does can be worse than causing bodily harm directly. Details matter; for instance, if it knocks out electricity and the refrigeration that is necessary for the protection of the food supply, a modest cyberattack could cause starvation.

ILLUSTRATION BY GARY NEILL

Proportionality

Proportionality in just-war theory is the idea that it would be wrong to cause more harm in defending against an attack than the harm of the attack in the first place; this idea comes from utilitarian ethics and is also linked to the notion of fairness in war. For example, a cyberattack that causes little harm should not be answered by a conventional attack that kills hundreds.[3,13] As one U.S. official described the nation's cyberstrategy, "If you shut down our power grid, maybe we will put a missile down one of your smokestacks."[5]

A challenge to proportionality is that certain cyberattacks, like viruses, might spiral out of control regardless of the attackers' intentions. While those consequences could be tolerated to prevent even worse consequences, lack of control means an attack might not be able to be called off after the victim surrenders, violating another key law of war. Another issue is that the target of a cyberattack may have difficulty assessing how much damage they have received. A single malfunction in software can cause widely varied symptoms; thus victims may think they have been damaged more

than they really have, and counterattack disproportionately. Therefore, counterattack—a key deterrent to unprovoked attacks—is now fraught with ethical difficulties.

Attribution

Discrimination in just-war theory also requires that combatants be identifiable to clarify legitimate targets—the principle of attribution of attackers and defenders. Terrorism ignores this requirement and therefore elicits moral condemnation. A problem with cyberwarfare is that it is very easy to mask the identities of combatants[4] Then counterattack risks hurting innocent victims. For example, the lack of attribution of Stuxnet raises ethical concerns because it removed the ability of Iran to counterattack, encouraging Iran toward ever more extreme behavior.

Attribution is an issue not only of moral responsibility but also of criminal (or civil) liability: we need to know who to blame and, conversely, who can be absolved of blame. To make attribution work, we need international agreements. We first could agree that cyberattacks should carry a digital signature of the attacking organization. Signatures are easy to compute, and their presence can be concealed with the techniques of steganography, so there are no particular technical obstacles to using them. Countries could also agree to use networking protocols, such as IPv6, that make attribution easier, and they could cooperate better on international network monitoring to trace sources of attacks. Economic incentives such as the threat of trade sanctions can make such agreements desirable.

Treacherous Deceit

Perfidy, or deception that abuses the necessary trust for the fair conduct of warfare, is prohibited by both Hague and Geneva Conventions. For instance, soldiers are not permitted to impersonate Red Cross workers and adversary soldiers. However, some ruses, misinformation, false operations, camouflage, and ambush of combatants are permissible. Cyberattacks almost inevitably involve an element of deception to make operations of a computer or network appear to be normal when they are not, as with tricking a user to click on a malicious link. So, to what extent could cyberattacks count as perfidy and therefore be all illegal given international humanitarian law?[9]

The moral impermissibility of perfidy is tied to the concept of treachery, and a prototypical example of a treacherous (and illegal) act in war is to kill with poison. Yet there are poisons that can kill quickly and painlessly, much more humanly than a bullet. This apparent paradox suggests the concept of treachery (and therefore perfidy) is fuzzy and difficult to apply. We do not get as angry when software betrays us as when people betray us. But maybe we should—software would be better if users were less complacent.

A Lasting Peace

In just-war theory, recent attention has focused on the cessation of hostilities and establishment of a lasting peace due to issues in recent insurgencies.[7] The consensus is that combatants have obligations after the conflict is over. For example, an attacking force might be obligated to provide police forces until the attacked state can stabilize, or attackers might have duties to rebuild the damage done by their weaponry.

This suggests that cyberattacks could be morally superior to traditional attacks insofar as they could be engineered to be reversible.[9,11,12] When damage done is to data or programs, the originals may be restorable exactly from backup copies, something that has no analogy with guns and bombs. Clever attacks could even use encryption to make reversal a decryption. Such restoration could be done quickly if the attack was narrowly targeted, and could be done remotely, so mandating reversal of cyberattacks after hostilities have ceased by the attacker could even become part of the laws of war. However, reversibility is not guaranteed when it is unclear what is damaged or so much is damaged that restoration takes an unacceptable amount of time. We need to establish ethical norms for reversibility and make them design requirements for cyberattack methods.

Insofar as most cyberattacks do not directly target lives, are they as serious?

Conclusion

The issues we have outlined in this column are only some of the basic ethical questions we must resolve if national cyberpolicies are to be supported by consistent and effective principles. And the right time to investigate them is prior to the use of cyberweapons, not during an emotional and desperate conflict or only after international outcry. By building ethics into the design and use of cyberweapons, we can help ensure war is not more cruel than it already is. C

References

1. Allhoff, F. *Terrorism, Ticking Time-Bombs, and Torture.* University of Chicago Press, In press.
2. Aquinas, T. *Summa Theologica.* Translated by Fathers of the English Dominican Province. Benziger Books, New York, 1948.
3. Coady, C.A.J. Terrorism, morality, and supreme emergency. *Ethics 114* (2004), 772–789.
4. Dipert, R. Preventive war and the epistemological dimension of the morality of war. *Journal of Military Ethics 5*, 1 (2006), 32–54.
5. Gorman, S. and J. Barnes. Cyber combat: Act of war: Pentagon sets stage for U.S. to respond to computer sabotage with military force. *Wall Street Journal* (May 31, 2011); http://online.wsj.com/article/SB10001424052702304563104576355623135782718.html.
6. McMahan, J. *Killing in War.* Oxford University Press, U.K., 2009.
7. Orend, B. War. In *Stanford Encyclopedia of Philosophy.* Stanford University, 2005; http://plato.stanford.edu/entries/war/
8. Owens, W., Dam, K., and Lin, H., Eds. *Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities.* The National Academies Press, Washington, D.C., 2009; http://www.nap.edu/catalog.php?record_id=12651.
9. Rowe, N. The ethics of cyberweapons in warfare. International *Journal of Cyberethics 1*, 1 (2009), 20–31.
10. Schneier, B. The story behind the Stuxnet virus. *Forbes* (Oct. 7, 2010).
11. The ethics of cyberwarfare. *Journal of Military Ethics* 9, 4 (2010), 384–410.
12. Towards reversible cyberattacks. In *Leading Issues in Information Warfare and Security Research*, J. Ryan, Ed. Academic Publishing, Reading, U.K., 2011, 145–158.
13. Walzer, M. *Just and Unjust Wars: A Moral Argument with Historical Illustrations.* Basic Books, New York, 1977.

Patrick Lin (palin@calpoly.edu) is an assistant professor in the philosophy department at California Polytechnic State University, San Luis Obispo.

Fritz Allhoff (fritz.allhoff@wmich.edu) is an associate professor in the philosophy department at Western Michigan University, Kalamazoo, MI.

Neil C. Rowe (ncrowe@nps.edu) is a professor in the computer science department at the U.S. Naval Postgraduate School, Monterey, CA.