Prospectus

Literature Review: The Intersection of Cybersecurity and Machine Learning
(Technical Topic)

Ethical Issues Stemming from the Stockpiling of Zero-Day Vulnerabilities
(STS Topic)

By

Derek Johnson

April 4th, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____Derek Johnson_____

Technical Advisor: _____Rosanne Vrugtman_____

STS Advisor: _____Richard Jacques_____

Intro:

How should we think about cybersecurity? This is a question that society is going to be forced to answer. Over the last 80 years, America has had to wrestle with the implications of being the world's most powerful military. There is currently a battle being fought over who will become the most dominant power in cyberspace. We know that the US spends billions of dollars on cybersecurity and their spending is likely matched by other countries(Sanger et al., 2020, paras. 1-2). China, North Korea, and Russia have all been linked to major domestic cyber-attacks in the last 5 years(Eichensehr, 2017, p. 530). It is unclear what the U.S.'s long-term tactics for dealing with these threats will be but currently, there is no system for regulating the cyber assets currently used. These assets, which can take the form of vulnerabilities in software available to the public, do not have to be disclosed to the companies who created the software. In my STS thesis, I will investigate the ethical responsibility government agencies have to disclose the threats they discover. In addition to discussing cyber threats on a macro-scale, it is important to look at the technologies that play a role in identifying and protecting against them. Machine learning (ML) is a technological framework that has been successfully applied to many challenging computing problems. In my technical thesis, I will explore how this technology can be applied to protect networks.

Technical:

Machine learning is not a new field. In fact, since the dawn of computing, researchers have attempted to simulate the critical thinking of human beings. In his 1950 paper *Computing Machinery and Intelligence*, Allan Turing outlined the framework that all successive computers

would use. He described his Turing machines as being able to "carry out any operation which could be done by a human computer"(Turing, 1950, p. 435). However, in the last couple of years, due to increased access to computational resources and new frameworks, machine learning has been applied to more areas than ever before. While some of these areas, such as computer vision and autonomous driving have proved to be fertile ground for ML, cybersecurity has not seen the technology take hold as readily. Endpoint protection is the process of identifying potential risks on network entry points such as desktops and mobile devices(*What Is Endpoint Security?*, n.d., para. 2). I think that this area of cybersecurity has the potential to be a good application for ML as it involves analyzing large datasets. For my technical thesis, I will perform a literature study on endpoint protection using machine learning. Through this process, I will identify the work that is currently being done in this space, identify areas for future research, and try to determine what challenges face researchers attempting to apply machine learning in this area.

Currently, endpoint security is done primarily using Antivirus software. The problem with this approach, as identified by Raff (2017), is that antivirus relies on identifying patterns it has seen before but is not able to flag new malware even if it uses the same functionality (p. 1). This means that current antivirus software can be a brittle defense. Machine learning techniques have been applied by researchers to malware identification with great success. These techniques break down into two categories: those that can be extracted from bytecode and those that run in a "sandbox" to identify malware(Le et al., 2018, p. 2). I will perform an analysis of the efficacy of these two categories and explain how they are currently being implemented.

The biggest challenge to applying ML to cybersecurity comes from its black-box nature. The autonomous capabilities, while offering increased scalability, may also be a threat. The

absence of human supervision may facilitate attackers to use techniques that are custom designed to take advantage of blind spots in the ML system(Apruzzese et al., 2018). This is the major drawback to using ML in cybersecurity. Another challenge facing ML researchers is gathering data to train their models on. Using a data sample that overemphasizes one risk type can result in models that do the same. This is a major issue and there are researchers working to create comprehensive data sets and testing systems for new malware classification systems(Pendlebury et al., 2018, p. 12).

The purpose of this technical thesis is to create an easy introduction to the current research in machine learning applied to endpoint protection. Hopefully, this will prompt those with ML experience to apply their knowledge to the field of cybersecurity. Additionally, I would hope people with cybersecurity backgrounds will feel empowered to use the tools of ML in their own work.

STS:

Cyber assets have become weapons of mass destruction. In May of 2017, the WannaCry ransomware attack was launched by a group of hackers backed by the North Korean government. It was one of the most high-profile and damaging cyberattacks in history, affecting more than 200,000 computers and resulting in millions of dollars in damages. The system, which targeted older versions of Microsoft operating systems, had been developed by the United States National Security Agency several years prior but had not been disclosed to the vendor. It had been stolen and leaked by a group of international hackers. Through the efforts of security researchers around the world, Microsoft was able to ship emergency patches to their systems, but substantial harm had already been done. In the aftermath of the attack, Microsoft released a statement saying

“ we need governments to consider the damage to civilians that comes from hoarding these vulnerabilities and the use of these exploits”(*The Need for Urgent Collective Action to Keep People Safe Online*, 2017, para. 3).

Clearly, to the leadership at Microsoft, the WannaCry attack was an example of governmental overreach. They, and many others, called for reform in the wake of the attack (Kesan & Hayes, 2016, p. 755). It became clear that there was a lack of faith in government agencies to discern when to disclose exploits they had discovered. However, the fact remains that these assets were stockpiled for a reason. If the NSA were to alert software vendors of every security vulnerability they find, they would be limiting their own ability to perform offensive cyber-attacks. Wicker describes the choices presented to security agencies through the familiar lens of the trolly problem. By not disclosing software vulnerabilities, the agencies can use the unpatched attack vector to invade the infrastructure of our nation’s enemies, keeping us safe. However, with this option, it is possible for other actors to discover the same vulnerability and use it to further their own agendas. By disclosing vulnerabilities to software vendors, the security agencies ensure it will be harder for other actors to use this vulnerability in the future. This option does not give the agencies the strategic advantage they might have had if the vulnerability had not been disclosed. This is the core of the issue that I hope to discuss in my STS prospectus.

In my research on this topic, I have come across many interesting framings of the problem. Most of the viewpoints I have read on the issue of vulnerability disclosure can be divided into two ethical camps. The two views I have come across can be described as consequentialist and non-consequentialist.

The consequentialists argue that the optimal decision is the one that results in the most overall happiness. Thus, there is not an answer to the question of to stockpile or not to stockpile.

The voices in this camp tend to believe that the way to determine whether a vulnerability should be disclosed is to weigh its potential benefit against its potential danger. This normally involves the creation of a scoring system, such as the one proposed by Pell & Finocchiaro (2017, p. 1565).

The non-consequentialists believe that all vulnerabilities should be disclosed. In this way, they avoid any of the calculated risks taken by the utilitarian consequentialists. This argument is articulated by Wicker (2017) when he writes that the non-consequentialist approach "offers more traction, capturing our ethical intuition regarding the public risk that many think is inherent in zero-day exploits in particular and cyber warfare in general"(pg. 103).

In my analysis of these topics, I think that it will be important to contrast these two ethical frameworks. Another facet of this topic that I would like to explore is the way policy is currently being implemented and how this implementation could improve. Currently, when government agencies discover or purchase zero-day vulnerabilities they do not have a protocol for deciding when to disclose these vulnerabilities. During the Obama administration, the government created a system to try to analyze vulnerabilities called the Vulnerability Equities Process (VEP) but in its current form "there are no hard and fast rules governing the VEP"(Schwartz & Knake, 2016, p. 2).

Conclusion:

It is important for people to feel secure in cyberspace. It seems as though the trend toward digitizing all sensitive information will only continue. This poses important cybersecurity issues. It will be important to create robust systems that can detect dangerous files. Currently, the onus is on individuals to have good cybersecurity practice but hopefully, with the help of ML-

powered endpoint security, we can protect those most vulnerable to cyber-attacks. By explaining the current state of the technology, it may inspire individuals to consider working on problems they had not considered before. Going forward, it will also be important to consider the ethical implications of our national cybersecurity policies. This is an area that many people are not aware of. By drawing attention to the ethical issues facing security organizations, I hope to spark debate about the types of disclosure policies that should be implemented.

References

Ablon, L., & Bogart, A. (2017). *Zero days, thousands of nights: The life and times of zero-day vulnerabilities and their exploits*. RAND Corporation.

Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018). On the effectiveness of machine and deep learning for cyber security. *2018 10th International Conference on Cyber Conflict (CyCon)*, 371–390. https://doi.org/10.23919/CYCON.2018.8405026

Boukerche, A., & Coutinho, R. W. L. (2021). Design Guidelines for Machine Learning-based Cybersecurity in Internet of Things. *IEEE Network*, *35*(1), 393–399. https://doi.org/10.1109/MNET.011.2000396

Eichensehr, K. (2017). Public-Private Cybersecurity. *Texas Law Review*, *95*(3), 467–538.

Full Text PDF. (n.d.). Retrieved April 5, 2021, from https://arxiv.org/pdf/1710.09435

Hausken, K., & Welburn, J. W. (2020). Attack and Defense Strategies in Cyber War Involving Production and Stockpiling of Zero-Day Cyber Exploits. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-020-10054-z

Kesan, J., & Hayes, C. (2016). BUGS IN THE MARKET: CREATING A LEGITIMATE, TRANSPARENT, AND VENDOR-FOCUSED MARKET FOR SOFTWARE VULNERABILITIES. *Arizona Law Review*, *58*(3), 753–830.

Le, Q., Boydell, O., Mac Namee, B., & Scanlon, M. (2018). *Deep learning at the shallow end: Malware classification for non-domain experts*. https://doi.org/10.1016/j.diin.2018.04.024

Pell, S., & Finocchiaro, J. (2017). The Ethical Imperative for a Vulnerability Equities Process and How the Common Vulnerability Scoring System Can Aid that Process. *Connecticut Law Review*, *49*(5), 1549–1589.

Pendlebury, F., Pierazzi, F., Jordaney, R., Kinder, J., & Cavallaro, L. (2018). TESSERACT: Eliminating Experimental Bias in Malware Classification across Space and Time. *ArXiv:1807.07838 [Cs]*. http://arxiv.org/abs/1807.07838

Raff, E., Barker, J., Sylvester, J., Brandon, R., Catanzaro, B., & Nicholas, C. (2017). *Malware Detection by Eating a Whole EXE*. https://arxiv.org/abs/1710.09435v1

Sanger, D. E., Perlroth, N., & Barnes, J. E. (2020, December 16). Billions Spent on U.S. Defenses Failed to Detect Giant Russian Hack. *The New York Times*. https://www.nytimes.com/2020/12/16/us/politics/russia-hack-putin-trump-biden.html

Schwartz, A., & Knake, R. (2016). *Government's Role in Vulnerability Disclosure: Creating a Permanent and Accountable Vulnerability Equities Process*. Belfer Center for Science and International Affairs. http://search.proquest.com/policyfile/docview/1923919527/CA7AE66278474220PQ/3

Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Wang, G., Welburn, J. W., & Hausken, K. (2020). A Two-Period Game Theoretic Model of Zero-Day Attacks with Stockpiling. *Games*, *11*(4), 64. https://doi.org/10.3390/g11040064

What Is Endpoint Security? | McAfee. (n.d.). Retrieved April 5, 2021, from https://www.mcafee.com/enterprise/en-us/security-awareness/endpoint.html

Wicker, S. B. (2021). The ethics of zero-day exploits---: The NSA meets the trolley car. *Communications of the ACM*, *64*(1), 97–103. https://doi.org/10.1145/3393670

Wolf, M. J., & Fresco, N. (2016). Ethics of the software vulnerabilities and exploits market. *The Information Society*, *32*(4), 269–279. https://doi.org/10.1080/01972243.2016.1177764

Yinka-Banjo, C., & Ugot, O.-A. (2020). A review of generative adversarial networks and its application in cybersecurity. *Artificial Intelligence Review*, *53*(3), 1721–1736. https://doi.org/10.1007/s10462-019-09717-4