Undergraduate Thesis Prospectus

Testing BLUESPAWN

(technical research project in Computer Science)

Organizing for Privacy: How Data Privacy Advocates Advance Their Agenda

(sociotechnical research project)

by

Will Mayes

November 1, 2021

technical project collaborators:

Calvin Krist

James McDowell

Jake Smith

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.

Will Mayes

Technical advisor: Yonghwi Kwon, Department of Computer Science

STS advisor: Peter Norton, Department of Engineering and Society

General Research Problem

How can online data privacy and integrity be protected?

Internet users generally value privacy, but companies value users' data, collecting it to target advertising (Martin & Murphy, 2016). Many internet users report concerns about privacy while "huge availability of data attracts abusive usage" (Schomakers et al., 2020). Cybercriminals also value and strive to obtain online user data (Mapmile & Mangoale, 2019). Identity theft is an extreme invasion of personal privacy. Because both legal and illegal data collection can constitute invasions of digital privacy, both regulatory and criminal enforcement responses are necessary.

Testing BLUESPAWN

How can the efficacy of BLUESPAWN, an open-source anti-malware tool, be measured and ensured?

I'm working on this project in the CS department under Yonghwi Kwon. The goals are to add testing to BLUESPAWN to measure and record current performance and ensure proper performance in the future. The solution must be accepted by other developers of BLUESPAWN.

BLUESPAWN is an open-source, active defense tool for Windows computers (Smith & Krist, 2020). The project's goal is to provide a tool for cybersecurity professionals to respond to advanced cyber threats. Endpoint Detection & Response (EDR) platforms are the state-of-the-art antivirus tools, but come with the downside that they function largely as black boxes. BLUESPAWN provides insight into detections by identifying malware behaviors as defined by the MITRE (2019) ATT&CK Framework (McDowell, 2021). BLUESPAWN is gaining attention from the cybersecurity community, having received over 800 stars on Github, but is not yet ready

for at scale use. BLUESPAWN can be found at https://github.com/ION28/BLUESPAWN (Smith et al, 2020).

For any software to be usable at large scale, developers must extensively test it. BLUESPAWN has yet to include such tests. By designing and developing a full suite of tests for BLUESPAWN, I will aid the project in becoming reliable at scale.

Current software testing is primarily done through unit testing, where developers test parts of a software individually. Software developers also use system tests, where a software's effect on a system is measured as a whole. This system can be as small as the software itself, and as large as the corporate network the software runs on. Finally, more advanced testing can include "fuzzing". Fuzzing involves generating random test cases to run against the program (Wang & Kang, 2018), often to look for crashes.

I will apply all three of these methods in testing. System testing will measure effects on the operating system by applying malware techniques to a Windows computer and verifying that BLUESPAWN identifies and remediates the issues. Fuzzing will be done using WinAFL, a windows specific implementation of American Fuzzy Lop (Zalewski, 2017), to catch crashes within BLUESPAWN. Unit tests will use Microsoft's Unit Testing Framework for C++.

At the end of this project, I will create a complete set of system and unit tests to verify accuracy of the current and all future builds of BLUESPAWN. I will also create a system for more tests to be added as necessary. I will ensure that all crashing inputs to BLUESPAWN have been identified and remediated through fuzzing, and the testing will be repeatable on future builds. This will ensure that BLUESPAWN is a reliable and accurate tool for catching malware and will remain so as development continues.

Organizing for Privacy: How data privacy advocates advance their agenda

How do digital privacy advocacy groups advance their agendas?

While online user data is collected and monetized on a vast scale (Martin & Murphy, 2016), privacy advocates seek to limit the practice.

The Electronic Frontier Foundation (EFF, 2021) works to hold companies accountable for consumer privacy. In September 2021, it organized a protest against Apple software that it contends "will endanger the privacy … of its customers." By then, privacy advocates had already compelled Apple to delay releasing the software. Apple cited concerns from "customers, advocacy groups, researchers, and others."(Apple, 2021) Privacy International (2020) uses ad targeting campaigns to expose "how difficult it is to understand how our data's used." Fight for the Future (2020) lobbies to ban facial recognition that "poses a threat to human society and basic liberty." Californians for Consumer Privacy (2020, November 4) ran a political campaign to pass California Prop 24 to "give Californians the strongest online privacy rights in the world." This law mandates that businesses obtain additional consumer permissions for collection of sensitive personal information. This law was passed in November, 2020 despite opposition from No On Prop 24 (2020) and other groups. To empower users, The Tor Project (2020) seeks to develop "censorship circumvention tools."

In 2020, California's Lieutenant Governor, Eleni Kounalakis, joined Californians for Consumer Privacy (2020, November 1) an interest group calling for the passage of Proposition 24.

Martin and Murphy (2016) argued that digital privacy scholarship is young and must be broadened. Introna and Gibbons (2009) reason that online advocacy groups serve as a shield against state surveillance, but they tend to focus on western countries. Bernstein and Hoffman

(2018) propose a framework for evaluating relationships between subnational and national coalition groups with regards to decarbonization. Their framework may be useful in the study of online privacy advocacies.

References

Apple. (2021, Sep. 3). Expanded Protections for Children. Apple.com.

Bernstein, S., & Hoffmann, M. (2018). The politics of decarbonization and the catalytic impact of subnational climate experiments. *Policy Sciences, 51(2), 189–211*. Web of Science.

Californians for Consumer Privacy. (2020, November 4). *California voters approve Prop 24*. Yes on Prop 24. https://www.caprivacy.org/california-voters-decisively-approve-prop-24/.

Californians for Consumer Privacy. (2020, November 1). (*Consumer and privacy advocates join California leaders in support of Prop 24*. Yes on Prop 24. https://www.caprivacy.org/consumer-and-privacy-advocates-support-of-prop-24/.

EFF. (2021, Sep. 9). Electronic Frontier Foundation. EFF Activists To Lead Protest Demanding Apple Cancel iPhone Scanning Program and Keep Its Privacy Promises To Customers. https://www.eff.org/press/releases/eff-activists-lead-protest-demanding-apple-cancel-iphone-scanning-program-and-keep.

Fight for the Future. (2020, June 18). Ban Facial Recognition. *https://www.banfacialrecognition.com/*

Introna, LD. & Gibbons, A. (2009). Networks and Resistance: Investigating online advocacy networks as a modality for resisting state surveillance. *Surveillance & Society. 6(3), 233-258*. Web of Science.

Mapimele, F., & Mangoale, B. (2019). The Cybercrime Combating Platform. *International Conference on Cyber Warfare and Security 2019*, 237–242.

Martin, K. D., & Murphy, P. E. (2016). The role of data privacy in marketing. *Journal of the Academy of Marketing Science, 45(2), 135–155*. Web of Science.

McDowell, J. (2021, May 17). *BLUESPAWN Design and Architecture*.

MITRE Corporation. 2019. Matrix - Enterprise | MITRE ATT&CK®. MITRE ATT&CK. *https://attack.mitre.org/matrices/enterprise/windows/*

No On Prop 24. (2020, September). *Los Angeles Area Chamber of Commerce Opposes Prop 24: Joins Consumer, Privacy and Small Businesss Advocates*. No On 24 CA. https://noon24ca.org/wp-content/uploads/2020/09/No-on-24-Release.LA-Chamber.pdf.

Privacy International. (2020, Sep. 24). Advertisers on Facebook: who the heck are you and how did you get my data? *https://privacyinternational.org/campaigns/advertisers-facebook-who-heck-are-you-and-how-did-you-get-my-data*

Schomakers, E.-M., Lidynia, C., & Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. *Electronic Markets, 30(3), 649–665*. Web of Science.

Smith, J., Krist, C. (2020, May 8). *BLUESPAWN: An Open-Source, Active Defense & Endpoint Detection and Response (EDR) Software for Windows-based Systems*.

Smith, J., Krist, C., Mayes, W., & McDowell, J. 2020. BLUESPAWN. *https://github.com/ION28/BLUESPAWN*

The Tor Project. (2020). The Tor Project Annual Report 2019-2020. The Tor Project.

Wang, C., & Kang, S. (2018). ADFL: An Improved Algorithm for American Fuzzy Lop in Fuzz Testing. In *Cloud computing and security: 4th International Conference, ICCCS 2018, Haikou, China, June 8-10, 2018: Revised selected papers. Part V* (pp. 27–36). essay, Springer.

Zalewski, M. (2017). *American fuzzy lop (2.52b)*. american fuzzy lop. https://lcamtuf.coredump.cx/afl/.