IMPE: Intelligence Malware Processing Engine

Feasibility of Different Levels of Vehicular Automation

A Research Paper
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degrees
Bachelor of Science in Computer Science and Computer Engineering

By
Dennis Tian

December 1, 2023

Technical Team Members:
None

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.

ADVISORS

Joshua Earle, Department of Engineering and Society

Briana Morrison, Department of Computer Science

Adam Barnes, Department of Electrical and Computer Engineering

Introduction

Autonomous driving has been a dream for many since as early as the 1930s. However, it has only recently become more concrete and feasible, with corporations like Google and Tesla leading the way (Reiser, 2021). It has the potential to save countless lives every year and resolve the problem of traffic, so it has slowly become a top priority of many automotive companies. However, the path to the adoption of this technology is far from straightforward. There are many problems related to ethics, safety, and public policy that will need to be addressed in the release of any form of autonomy (Fleetwood, 2017). Furthermore, there are different levels of autonomy, ranging from simple lane assist, which already exists commercially, all the way to full autonomy. The key question that will be investigated in my research will be how policy, regulation, and technical hurdles have shaped the industry's push toward full autonomy.

My technical project is not related to my STS project; it describes the experience I had working as a summer intern on a cybersecurity application development team for the past two summers. I was part of a team in charge of a malware analysis web application, where security analysts were able to feed in a sample and receive a report containing information such as extracted strings, malware family, dropped payloads, and malicious behavior. The main goal of this application was to speed up the process of malware analysis and incident response, automating many of the tedious processes needed to triage a suspected file.

I will first elaborate on the technical project, providing technical and implementation details. Then, I will delve into the STS project; I will first give some more background, then define the relevant social groups I will focus on, explain the STS framework I will use, and explain the methods I will employ during this research. The prospectus will conclude with a brief review of some foundational and significant texts that may be valuable for subsequent research.

Technical Project

Security analyst and incident response (IR) teams play a key role in preventing, detecting, and mitigating cybersecurity attacks; however, many of the necessary steps are repetitive and time consuming. The team I worked with for the past two summers was in charge of the development of an internal web application, the Intelligence Malware Processing Engine (IMPE), that enables submission of a piece of suspected malware and returns valuable metadata and behavioral information obtained from various static and dynamic analysis tools. The project was divided into two main teams: those in charge of the web application itself and those overseeing the reverse engineering (RE) tools that the application uses to process samples.

The web application side of IMPE was written using mostly Python 3 and JavaScript and had extended support for API integrators, people who interact with the application through the REST API rather than the UI. We implemented IMPE's backend using a combination of Celery, RabbitMQ, and Redis for parallelized worker/queue management, and we wrote the frontend using the React framework. The RE tools were written in a variety of languages, including Python, YARA, and Rust. They encompass both static and dynamic analysis and include both third-party programs and ones written in-house.

The tools I worked on mostly focused on static file analysis, where the goal was to glean information from a file without actually running it. For example, I helped design and write parsers to extract keywords and strings that could provide insight on the behavior of a malicious file. As another example, I wrote malware signatures using YARA; these signatures each correspond to a malware family and are run against files to potentially determine what family they belong to. IMPE has gradually become the go-to resource for analysts whenever they need information about a sample that would otherwise take hours or days to compile by hand. The

frontend and API are currently undergoing a major facelift (which I also worked on) that will allow it to be integrated into a larger “Analyst Tools” portal that contains a suite of other tools. The end goal of this portal is to have everything polished and tested to be ready for commercial distribution.

In addition to the core IMPE project and the RE tools, I also worked on several auxiliary features, with two notable ones being CronIngest and the IMPE Slackbot. CronIngest, as the name hints to, is a program that automatically ingests new samples from VirusTotal’s database using their Retrohunt feature. This feature allows CronIngest to query for samples in a time range that match on a provided set of rules, written in YARA. After the samples are obtained, they are then relayed over to IMPE for processing. This information is crucial to analysts because it allows them to keep an eye on new and upcoming vulnerabilities without having to dedicate the time for manual research. As mentioned, I also was the primary developer on the IMPE Slackbot. The primary form of communication at the company is Slack, so having the ability to use IMPE directly from Slack was an oft-requested feature. I was able to integrate the IMPE and Slack APIs to create an app that could respond to slash commands like `/impe search <hash>`. This bot helped to greatly increase the productivity of analysts, as switching back and forth between Slack and a web browser to access IMPE was both tedious and prone to errors when copying information.

STS Project

I will be investigating a technology that is becoming increasing familiar in everyday life - autonomous vehicles. Although the technology has definitely not manifested itself fully in current implementations, it is undoubtedly a goal that the industry is pushing toward. This is a relatively controversial topic in many realms, notably the in the legal and insurance departments. Obviously, the category of autonomous vehicles is extremely broad. However, there are numerous issues and societal effects that it has, and the research question I will be trying to answer is how safety, legal, and other concerns as well as major setbacks have influenced not only the progress of autonomous vehicle adoption but also the technology itself.

In principle, a large-scale system of fully autonomous vehicles can greatly reduce traffic congestion, fuel consumption, and accidents (Fleetwood, 2017). However, the road to this idealistic scenario has been and will be full of obstacles, both expected and unexpected.

The first concern that likely comes to mind is safety—having a black-boxed technology drive your vehicle while you divert attention from the road may understandably make some feel uneasy; this is exacerbated by the fact that the technology itself is not and will never be perfect. Companies on the forefront of autonomous vehicle innovation have dealt with numerous accidents, most recently being a gruesome incident where one of Cruise's robotaxis dragged a woman 20 feet after hitting her on a road (Marshall, 2023). Events like these tarnish the public perception of not only the company involved but the autonomous vehicle industry as a whole. These kinds of scenarios stem from the fact that software for autonomous vehicles is extraordinarily difficult to refine. There are endless possibilities of road conditions, and developing a model that can handle edge cases reliably is a near-impossible task. Even if the

technology were mature enough to release to the public, reception might not be warm due to lack of trust stemming from prior happenings.

The other major challenge that this technology has faced is policy and regulation (related in part to safety). One particular problem that development in this field has experienced is regarding testing. Simulated environments only go so far when trying to develop models that respond accurately to real-life scenarios; however, the dilemma of testing in non-laboratory circumstances is that the product is still potentially unfinished, posing danger to anyone present. The Cruise incident serves as a prime example of this: unable to continue testing in the United States for the time being, they pivoted to Japan and Dubai to continue testing while safety investigations occurred in the United States (Bensinger, 2023). This, of course, also poses ethical questions on how the technology and the testing of it should be regulated.

Autonomous vehicles have the potential to affect a significant number of people, even ones who do not use the technology. This will ultimately affect some social groups more than others, though. One category I have chosen to focus on are those who may be enabled to drive better or drive at all – specifically, the elderly and disabled. Another relevant social group I will be exploring are teenage and young adult drivers. These drivers are known to be more reckless on average and have a higher incident of driving under the influence (DUI) (Kasperowicz, n.d.), so it will be fascinating to investigate the effect of autonomy on them now and in the future. However, I will be leaving out what is probably the largest of these social groups: healthy, safe drivers. This is because for the most part, these people will be the least affected by autonomous vehicles, with it mostly just being a convenience rather than a core improvement to everyday life.

The STS framework I will be using for this paper will be actor-network theory, or ANT. This framework asserts that a technology and how it interacts in the social landscape can be represented as a network comprised of both human and non-human actors. These actors all have individual interests, and they may work with or against each other to shape the network (Sismondo, 2010). ANT is a suitable framework for analyzing my research question because it facilitates the examination of all of the different relationships between various facets of the technology, some of which may not have been initially apparent. This comprehensive analysis can then be used to identify actors and relationships that may be used to support the technology's development and ones that should be addressed to avoid setbacks.

Presently, I am planning to employ two primary methods to aid in research: finding, reading, and synthesizing previous literature and case studies. The first step in my research will be to analyze literature regarding the history of autonomous vehicles to gain a more comprehensive understanding of its evolution, both as a technology and a societal influence, as well as the different levels of autonomy. I will then investigate the current progress of autonomy, including aspects that are still in development; this will encompass the technologies themselves, the benefits they have been able to provide, and the difficulties that accompanied them. After that, I will review case studies centering on specific companies and incidents that impacted their progression regarding this technology. This will give insight on the impact autonomous vehicles have on society and how the public perceives them, and it may also shine light on recurring or common hurdles that have been encountered.

Significant Texts

The relationship between autonomous vehicles and society is a complex and controversial one; therefore, numerous journal articles and informational sites will need to be

consulted. One of the foundational texts that relate to this topic is an article published in the American Journal of Public Health, titled *Public Health, Ethics, and Autonomous Vehicles* and written by Janet Fleetwood. The text investigates the efficacy of autonomous vehicles from a public health and ethics standpoint, as the title suggests. Ultimately, it calls for public health leaders to take action with regards to autonomous vehicles, and for them to embrace autonomous vehicles as an instrument of traffic safety (Fleetwood, 2017). This is an important piece for my research because it provides specific numbers on how effective autonomous vehicles can be on traffic fatalities, funding, and more.

Another critical piece is Andreas Herrmann's book *Autonomous Driving: How the Driverless Revolution Will Change the World*. This book provides a wealth of information regarding almost all facets of the topic. However, the main section I will likely be using is the one that discusses and defines the five levels of vehicular autonomy (Herrmann, 2018). This will be necessary when analyzing the current progress of autonomy and how further development will affect society.

Besides the concepts of automated driving, I will also need information about the current generation of this technology, including advancements and challenges. Venkata Kosuru and Ashwin Venkitaraman's *Advancements and challenges in achieving fully autonomous self-driving vehicles* provides an in-depth analysis of these aspects. Notably, it states that current-day autonomous vehicles have not yet passed safety and road tests, but advancements are rapid and promising.

Greg Bensinger's article *Cruise testing continues in Japan, Dubai, even as vehicles parked in US* provides insight into the recent Cruise testing recall in the United States and how the company has pursued continued testing in Dubai and Japan. The recall was due to an

accident a month earlier. The article touches on a critical aspect in the process of autonomous vehicle adoption: testing. Even if a driver is present in an autonomous vehicle, there is still a risk of an accident. However, testing would likely be most effective in real-world scenarios, but these scenarios have the highest intrinsic danger. The article presents a specific incident that I can use to bolster my argument.

Finally, any decisions on the adopting of this technology will be rooted in the effects on society. This will clearly need to be investigated, and Hussain et al.'s article *Autonomous cars: Social and economic implications* will be an invaluable resource in this process. The article discusses risks the technology may pose in various domains, from direct users to the job market.

References

Bensinger, G. (2023, November 18). Cruise testing continues in Japan, Dubai, even as vehicles parked in US. *Reuters*. https://www.reuters.com/business/autos-transportation/cruise-testing-continues-japan-dubai-even-vehicles-parked-us-2023-11-17/

Fleetwood, J. (2017). Public health, ethics, and autonomous vehicles. *American Journal of Public Health*, *107*(4), 532–537. https://doi.org/10.2105/AJPH.2016.303628

Herrmann, A., Brenner, W., & Stadler, R. (2018). *Autonomous driving: How the driverless revolution will change the world*. Emerald Publishing Limited. https://doi.org/10.1108/9781787148338

Hussain, R., Lee, J., & Zeadally, S. (2018). Autonomous cars: Social and economic implications. *IT Professional*, *20*(6), 70–77. https://doi.org/10.1109/MITP.2018.2876922

Kasperowicz, L. (n.d.). Teens or seniors: *Which group causes more accidents?* Insurance.Com. Retrieved October 25, 2023, from https://www.insurance.com/auto-insurance/safety/teens-or-seniors-who-are-our-worst-drivers.aspx

Marshall, A. (2023, October 24). *GM's cruise loses its self-driving license in San Francisco after a robotaxi dragged a person*. Wired. https://www.wired.com/story/cruise-robotaxi-self-driving-permit-revoked-california/

Reiser, A. (2021, August 9). History of autonomous cars. *TOMORROW'S WORLD TODAY®*. https://www.tomorrowsworldtoday.com/2021/08/09/history-of-autonomous-cars/

Sismondo, S. (2010). *An introduction to science and technology studies* (2nd ed). Wiley-Blackwell.

Venkata Satya Rahul Kosuru & Ashwin Kavasseri Venkitaraman. (2023). Advancements and challenges in achieving fully autonomous self-driving vehicles. *World Journal of*

Advanced Research and Reviews, 18(1), 161–167.

https://doi.org/10.30574/wjarr.2023.18.1.0568