Machine Learning models to address security concerns with attacks in Autonomous Vehicle motion planning systems

Exploring the implications of Autonomous Vehicles on urban social inequality

A Thesis Prospectus
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 Degree
Bachelor of Science in Computer Engineering

By
Shrisha Yapalparvi

November 3, 2023

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

Prof. Pedro Augusto P. Francisco, Department of Engineering and Society

Prof. Adam Barnes, Department of Electrical and Computer Engineering

Introduction

The advent of general computing machines for the everyday consumer starting in the 1980s marked a pivotal point in human technological history. (University of Rhode Island, n.d), Since then, both dependence on and research / development of computer technologies has grown exponentially. The latest frontier of computing technology is coming in the form of Artificial Intelligence (AI) and robotics (Helfrich, 2022). At the intersection of these two technologies is the pursuit of autonomous systems. While such systems have been used for commercial applications, such as manufacturing and logistics services, for decades, only in the past couple of years have autonomous systems become something that the public can interact with. Among the most notable examples of autonomous systems is Autonomous Vehicles (AVs). AVs can take the form of personal cars, rail lines, trucks, planes, and more, all with the promise of revolutionizing the way humans and goods travel (Litman, 2023).

The focus of this paper will be on autonomous personal cars. The development of AV systems includes the integration of many exciting, cutting-edge technologies including AI, sensors, and embedded communication protocols. These technologies work together to collect data form the external world and process them into movements. However, due to the ever-growing complexity of these systems, specifically the AI systems used to help plan the movements of the car, it becomes difficult to implement adequate defenses against attacks (Mohseni et al., 2019). Despite multiple levels of backup and redundancy for the hardware and mechanical components of AVs, little has been done to ensure the proper performance of models under adverse conditions and attacks.

As with any emerging technology, it is also important to consider the social impact before its widespread adoption. AVs are a rapidly developing technology that promise to revolutionize the way humans move, but very little work has been done studying the impact of AVs on different socio-economic groups. Existing studies have either not considered a diverse set of demographics, or, of the 20% that have, insert other biases in their methods of study. The exploration AV's social impact is beneficial to the public, the government, and the AV companies themselves (Bills, 2020). This information can be used to make sure that society harnesses the advantages of AVs without the pitfalls of exacerbating social inequality, an issue that becomes even more pressing considering the mass adoption of AVs is coming within the next 10 to 20 years (Litman, 2023).

With this context, the research will explore autonomous systems, both from a technical and social point of view. The topic of the technical research will explore the use of AI / Machine Learning (ML) models to monitor motion planning systems in AVs and apply patches as they are needed to keep the systems running securely. Relatedly, the socio-technical topic of this research will be regarding the effect of AVs (specifically, autonomous cars) on urban social inequality. Research in both domains would allow for rapid innovation in AV safety while promoting best practices and standards that can make the technology accessible equitable for all.

ML models to address security concerns with attacks in AV motion planning systems – Technical Topic

AVs are extraordinarily intricate systems that incorporate hundreds of different technologies to enable their autonomous operation. To understand how they work under the hood, its easiest to break it down into three main components: perception, motion planning, and control (Litman, 2023). The perception stack is responsible for getting data from the outside world so that the computer can process it. This includes sensors such as cameras that can take in visual data, inertial measurement units and global positioning systems that take in positional data, and LIDAR / RADAR, which use light rays and radio waves respectively to create a map of objects around them. The data from these sensors are combined in a process known as sensor fusion to simultaneously localize and map the vehicle so it knows where it is, and where everything around it is (Kocić et al., 2018). The motion planning stack takes this data and passes it through an algorithm to determine the future motion of the car. These models can vary from being algorithmic (mathematical equations based on a distance to a wall, for example) to being built with ML / AI that uses previous data to identify what steps it needs to take to get it to the state it wants to be in. Finally, control takes the decisions made by the motion planning stack and converts it to movements in the real world. Using actuators and embedded communication protocols, vehicles can turn, brake, and accelerate (Valigi, 2018).

Most modern-day motion planning stacks are built using AI, specifically Neural Networks (NNs). NNs are a special type of AI architecture that mimics the way the human brain learns. Convolutional NNs (CNNs) are used due to their strength in image processing, object recognition, and object detection (Haija, Gharaibeh, & Odeh, 2022). On top of CNNs, Long Short-Term Memory (LSTM) is also used, which is a NN architecture that uses previous data combined with present data (whereas traditional NNs only use present data) to make decisions. This gives the computer access to context, which helps it predict the motion of dynamic objects like other vehicles and pedestrians (Eraqi et al., 2017).

These NN models grow in complexity both with the number of internal nodes (which are set by the ML engineers), and with the amount of training data (Hardesty, 2017). Motion planning stacks, with access to terabytes of data and massive internal architecture can reach sizes of billions of parameters, which makes it extremely difficult for engineers to understand exactly what the model is optimizing for internally. This is where a massive security risk exists. Attacks on these models can happen by simply altering a couple pixels as the data is fed into the model. Most of the time, these changes are not noticeable to the human eye but can throw off predictions made by the motion planning stack significantly (Zhou et al., 2023). Most AV manufacturers account for and add redundancy to mechanical and hardware failures (brakes not working, for example), but little is being doing to address the safety performance of these AI models (Mohseni et al., 2019). Thus, an opportunity arises to explore the development of ML models that can act as a watchdog for the existing motion planning stack. These models can detect anomalies in both the data coming into the stack and the controls going out to increase AI model performance safety. The technical research will focus on the development of ML models to detect and mitigate attacks on the models used in AV motion planning stacks.

AVs and their effect on urban social inequality – Science, Technology, and Society Topic

It is undeniable that autonomous systems can bring massive benefits to humanity, and when looking at AVs specifically, the benefits are multifold. They can boost productivity by allowing humans to work, sleep, and enjoy entertainment instead of being tied up by driving. AVs can also expand access to personal transport, offering increased personal mobility to the 13% of American disabled / elderly people who could not drive themselves previously (Urban Institute, 2022). Traffic can be reduced by partitioning AVs such that they are always on the road transporting people and optimizing routes. This would also get rid of unwanted infrastructure such as parking lots, which takes away from valuable urban space (Wallace, 2017). AVs will also make transport cheaper, since most of the costs related to transport is labor. Studies show transport could become as cheap as 10 cents a mile (Litman, 2023). Perhaps the most important benefit is the increase in safety of road travel since the computer systems never from fatigue or influence of substances.

However, for every benefit that AVs provide, there are a list of challenges and obsticals they pose if not well regulated / deployed properly. Up to 1 in every 9 Americans have their career in some way related to the transport industry (including drivers, public transport workers, gas station workers, and motel workers), which means a significant portion of the labor force and thus, the economy, is at risk of job automation. AVs can also exacerbate wealthy suburbia, with wealthier people being able to move further away from cities and commute further distances while increasing productivity – leaving the less wealthy behind (Wallace, 2017). Most importantly, the quasi-public transit benefits would all be negated if those who owned / operated AVs decided not to share them. In fact, this type of behavior is already present with companies like Uber and Lyft circumventing public transit regulations – specifically requirements to become accessible – since they qualify themselves as tech companies instead (Urban Institute, 2022).

Governments across the world have already started asking themselves what they would like the future of AVs to look like. They want to see the benefits of AVs but are falling behind on the curve of adoption (Garrick & Atkinson-Palombo, 2019). There is a clear need for more social focused AV research. This is why the socio-technical research conducted is regarding the impact of AVs on urban social inequality. The research will focus on technological determinism perspectives and actor network theory regarding the AV technology itself, the manufactures, different socio-economic demographics, the government / policy makers, and the current transportation industry. Data can be collected how each of these actors exist and how they influence each other can help determine what aspects can cause future social inequality, and which aspects can help boost equity. This information can then be compiled to provide best practice guides to AV manufacturers, researchers, governments, and the public to ensure an equitable adoption of the technology.

Conclusion

As autonomous vehicles become a closer reality, it is important to consider both the social and technical implications of it so that when they hit the scene, it presents itself as the beneficial technology it promises to be. On the technical end, current work lacks proper redundancy for motion planning models (Mohseni et al., 2019). To address this, research will be done to create models that can detect attacks on motion planning ML systems as they happen and deploy patches to maintain the integrity of the vehicle. On the socio-technical end, little work has been done truly analyzing the impact of AVs on different socio-economic groups (Bills, 2020). To address this, actor network theory and technological determinism will be used to analyze how all stakeholders in the AV scene impact social inequality. The results of this work can be used to secure autonomous systems against attacks to make them safer while also giving companies and governments best practices to ensure all people reap the benefits of AVs.

References:

Bills, T (2020). On Transportation Equity Implications of Connected and Autonomous Vehicles (CAV) A Review of Methodologies. Final Report. USDOT CCAT Project No. 5.

Eraqi, H. M., Moustafa, M. N., & Honer, J. (2017). End-to-end deep learning for steering autonomous vehicles considering temporal dependencies. arXiv preprint arXiv:1710.03804.

Garrick, N., & Atkinson-Palombo, C. (2019). What do we want from autonomous vehicles (Avs)? Using participatory planning and scenario analysis of alternative features to identify stakeholders' desired outcomes from the strategic deployment of emerging transportation technology (Tech Report 2018 Project 12). University of North Carolina at Charlotte. Center for Advanced Multimodal Mobility Solutions and Education. https://rosap.ntl.bts.gov/view/dot/61501

Hardesty, L. (2017). Explained: Neural networks. MIT News. https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414

Helfrich, T. (2022, May). Council post: Why robotics and artificial intelligence are the future of mankind. Forbes. Retrieved November 1, 2023, from https://www.forbes.com/sites/forbestechcouncil/2022/05/31/why-robotics-and-artificial-intelligence-are-the-future-of-mankind/

History of computers. (2022). Retrieved November 1, 2023, from https://homepage.cs.uri.edu/faculty/wolfe/book/Readings/Reading03.htm

Kocić, J., Jovičić, N., & Drndarević, V. (2018, November). Sensors and sensor fusion in autonomous vehicles. In 2018 26th Telecommunications Forum (TELFOR) (pp. 420-425). IEEE.

Litman, T. (2023). Autonomous Vehicle Implementation Predictions Implications for Transport Planning. Victoria Transport Policy Institute. https://www.bilbloggen.dk/wp-content/uploads/2023/04/Autonomous-Vehicle-Implementation-Predictions.pdf

Mohseni, S., Pitale, M., Singh, V., & Wang, Z. (2019). Practical solutions for machine learning safety in autonomous vehicles. arXiv preprint arXiv:1912.09630.

Shared autonomous vehicles could improve transit access for people with disabilities if regulated appropriately | urban institute. (2022, October 4). https://www.urban.org/urban-wire/shared-autonomous-vehicles-could-improve-transit-access-people-disabilities-if-regulated

Valigi, N. (2021). Lessons learned building a self driving car on ros. In A. Koubaa (Ed.), Robot Operating System (ROS)(Vol. 895, pp. 127–155). Springer International Publishing. https://doi.org/10.1007/978-3-030-45956-7_5

Wallace, R. L. (2017). Mobility: The Socioeconomic Implications of Autonomous Vehicles. Science, Technology, and Public Policy Program, Ford School of Public Policy, University of Michigan. https://stpp.fordschool.umich.edu/sites/stpp/files/2021-07/Mobility%20The%20Socioeconomic%20Implications%20of%20Autonomous%20Vehicles.pdf

Zhou, X., Chen, A., Kouzel, M., Ren, H., McCarty, M., Nita-Rotaru, C., & Alemzadeh, H. (2023). Experimental security analysis of dnn-based adaptive cruise control under context-aware perception attacks. https://doi.org/10.48550/ARXIV.2307.08939