Threat Analysis for Cooperative Automated Driving

Sumeet Chhawri Derrick Dominic Ryan Eustice Di Ma André Weimerskirch

***Abstract*—Cybersecurity has become a necessary consideration for commercial automobile deployment as automotive technologies incorporate increasing complexity through electronics and connectivity. Now, as a result of global investment in Automated Driving (AD), security must be considered for these upcoming applications which transfer control from the human driver to the vehicle. While prior work in the field has assessed and started providing solutions to secure modern automobiles, there is much to understand regarding the security implications of AD. In this work, we propose and demonstrate a customizable risk assessment methodology for AD applications consisting of a generalizable reference architecture and a threat model. Our approach allows for the analysis of AD threats across all levels of automation leading to the identification of critical risks and respective potential solution strategies.**

I. Introduction

Automobiles have evolved into sophisticated machinery with the proliferation of Information and Communications Technology (ICT) with enhanced automotive safety and connectivity, through Advanced Driver Assistance Systems (ADAS) and Vehicle Communication (VC). More recently, another direction of this advancement is towards the realization of Automated Driving (AD). Since the Defense Advanced Research Projects Agency (DARPA) Grand and Urban Challenges, there has been a significant investment by both industry and academia to make AD a reality. Google, Ford, Volvo, and Mercedes among many others have shown credible results contributing towards autonomous vehicles which partially or entirely replace the human driver.

Apart from the complex algorithms and hardware requirements, the safety and privacy of future AD passengers will depend on the security solutions deployed to secure the vehicle from malicious and unintended cyber-attacks. The current space of automobile cybersecurity is immense, comprising physical and remote attack surfaces and industry specific challenges of longer product life of vehicles, longer development cycles to name a few. Koscher et. al. [1] introduced some of the initial work toward assessing automotive attack surfaces, showing that attackers with physical access to modern cars can exploit several internal vulnerabilities. This work was extended by Checkoway et. al. [2] who discussed remote attack surfaces in modern vehicles. In regards to AD, Petit and Shladover [3] have offered one of the earliest analyses of cybersecurity in automated and connected vehicles by identifying threats in full and high AD. The authors analyzed attack surfaces in sensors and infrastructure and incorporated cooperative AD through threat analysis of attack surfaces in Vehicle to Everything (V2X).

While prior work in automotive and AD systems addresses attack surfaces and approaches, the impact of potential threats is assessed without an understanding of how components work together to realize applications. However, different components are used in different ways to realize AD applications, and the way components are used may also be a source of vulnerabilities. Understanding the application under attack will additionally better inform about the effect of the attack for accurate risk analysis. We propose an application based approach to identification of threats. Our methodology uses an AD architecture as a reference to understand the critical paths of threats in an AD application. For risk assessment, we present an intricate and customizable threat model derived from National Highway Traffic Safety Administration (NHTSA) [4] and E-safety Vehicle Intrusion Protected Applications (EVITA) [5] incorporating threat actors and attack scenarios resulting in an enhanced threat matrix with tabular and visual representation of risks associated with attacks. Our approach can be used to provide an assessment of security concerns at different levels of automation to better understand which risks and their respective security solutions will need to be prioritized at different times.

II. Automated Driving Framework

Our approach to AD threat analysis begins at the AD application. Understanding how individual components work together to realize an application can better inform on the likelihood and impact of a particular attack.

A. Applications

While we do not exhaust the space of current or potential AD applications, we take note of a few representative applications, some of which we will explore in our threat analysis (Section IV). We categorize applications by their level of automation using the SAE definitions [6] and the AdaptIVe application classifications [7].

Level 0 (No Automation)—At Level 0, the human driver is completely responsible for the driving task. However, the driver may be passively assisted with warnings (e.g. for lane departure) or actively assisted in safety critical situations (e.g. forward collision emergency braking).

Level 1 (Driver Assistance)—A Level 1 AD system is able to execute exactly one of longitudinal or lateral control with all remaining functions performed by the human driver. Applications in this category include Adaptive Cruise Control (ACC) which only uses longitudinal control and Parking Assistance with Steering which only uses lateral control.

Level 2 (Partial Automation)—Where a Level 1 AD application can only execute one of longitudinal or lateral control, a Level 2 application can execute both. Applications include combined ACC and Lane Keeping Assistance (LKA), and Key Parking, a system which allows a driver outside the vehicle to

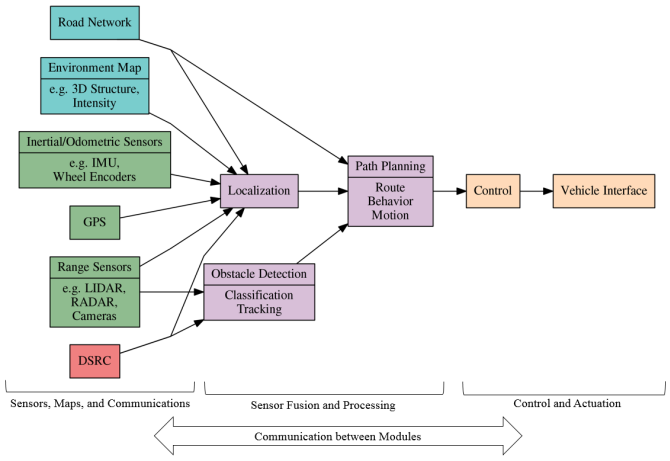

Fig. 1: An interconnection diagram of our generalized architecture for a AD system.

initiate and supervise an automated parking maneuver using a remote.

Level 3 (Conditional Automation)—In addition to controlling the vehicle actuators, applications at Level 3 can temporarily take over the responsibility of monitoring the vehicle and environment. However, if the vehicle and environment states leave the operational space of the application, the application can request human driver intervention. An application at this level would be the Traffic Jam Chauffeur, a system which performs the full driving task during traffic jams.

Level 4 (High Automation)—This is the final level where a human driver is still required to make decisions for the vehicle. Unlike a Level 3 system, a system at Level 4 must be able to direct the vehicle to a minimal risk condition before requiring the human driver to intervene. An example of such an application is Driverless Valet Parking, a system with which a human driver can exit the vehicle, initiate parking via a remote, and have the vehicle park itself in a garage or parking lot.

Level 5 (Full Automation)—At Level 5, the AD system is able to perform all aspects of the driving task without any human driver supervision. A representative application for this level is the fully autonomous taxi.

B. Reference Architecture

Given an AD application, we need to understand its implementation to determine potential attack surfaces and begin to enumerate threats. However, as most AD systems (particularly for high levels of automation) are still in development, there exists no standard architecture for a commercial AD solution. To that end, we provide a generalized architecture for an AD system that captures the state of the art and allows us to deconstruct applications at various levels of automation.

Our architecture is derived from research platforms developed for the DARPA Urban Challenge [8–10], further post-competition work by the Stanford Urban Challenge team [11], and the ETH V-Charge project [12]. An interconnection diagram for our architecture can be seen in Fig. 1.

Although this architecture is based on a fully automated system, we use functional blocks as a layer of abstraction above implementation. Each functional block conveys (at a high level) the information it acquires, how it processes that information, and what it produces as an output. This abstraction allows us not only to model systems with different implementation details but also to understand systems at lower levels of automation.

Our functional blocks consist of:

Sensors, Maps, and Communications—Sensors include inertial and odometric sensors for perceiving the vehicle state, a GPS receiver for global localization, and range sensors for perceiving the environment. Prior information about the environment is encoded in maps such as dense environment map for precise localization and more abstract road network maps for path planning. Vehicle to Infrastructure (V2I) and Vehicle to Vehicle (V2V) communication through DSRC can augment the internal sensors.

Sensor Fusion and Processing—Information from the sensors and maps is fused and processed in 3 modules. Localization is responsible for localizing the vehicle in its environment. Obstacle Detection detects, classifies, and tracks objects in the environment. Given the understanding of the vehicle and environment state from Localization and Obstacle Detection, Path Planning plans a trajectory following the rules of the road to achieve a high level goal.

Control and Actuation—The output of Path Planning is fed to the control module which executes the trajectory through the vehicle interface (steering, acceleration, and braking).

While our specification describes a standalone vehicle, we also consider elements outside of the vehicle and/or part of the infrastructure that comprise a commercial AD solution. These include but are not limited to Road Side Units (RSUs), servers for shared mapping, vehicle keys, personal electronics, and operational infrastructure for maintenance and diagnosis.

III. Threat Model

Threat modeling provides an identification of the threats, attacks and vulnerabilities existing within and to the system under consideration. There has been significant contribution towards threat modeling as applied to automobiles. We derive the proposed threat model for our work by combining the strengths and addressing the limitations of the NHTSA [4] and EVITA [5] threat models.

A. Limitations of Existing Threat Models

Threat modeling is a process which requires subjective evaluation of threats specific to the system under investigation. It is also an iterative process, with threat variables needing to be updated as new information becomes available. As they are, we feel that the NHTSA and EVITA approaches could be improved: the NHTSA approach fails to capture *Threat Actors* and has limited factors influencing *Motivation*, and the EVITA approach is specifically geared towards V2X based threats. While NHTSA and EVITA have contributed significantly to automotive threat modeling, this work serves as

a further iteration, considering new variables and introducing an enhanced visual depiction of the threat matrix.

B. Proposed Threat Model

The first step in creating the threat model is to identify and create and an exhaustive application list of the automobile product under consideration spanning the domains of safety, powertrain, body electronics, entertainment/infotainment functions, and categorizing based on the levels of AD, safety, operation and privacy. This allows a birds-eye-view of the possible applications an attacker may target for a successful cyber-attack. In tandem is to define the architecture under consideration. This allows the threat modeling team to understand the flow of information/data for the application under investigation and informs of vulnerabilities at each stage of information processing. For this work we have devised an AD framework as described in section II consisting of both the application list and reference architecture.

Once the application list and architecture are finalized, threat actors are investigated. Threat actors are the attacker profiles which describe an actor's capabilities and potential motivations. For the purpose and scope of this work we recommend the following threat actors: *Thiefs, Owners, Mechanics and service providers, Organized Crime, Hacktivists, Terrorists, National and international entities*. Conceptualization of an attack scenario derived from the application, AD framework and the threat actors with levels of abstraction is entailed, allowing us to define similar attacks with permutation of various actors, attack methods and different vulnerabilities in the AD framework. For example, *Car Theft* can be performed by Thiefs and Organized Crime. Attack scenarios can involve cloning the key fob or bypassing Anti-Theft systems and ignition through OBD-II port. The next variable in the model, the attack method, describes the low-level attack methodology implemented by the attacker. These are derived from the NHTSA threat model and include: *Spoofing, Tampering with Data, Repudiation, Information Disclosure, Elevation of Privilege and Malicious*.

Following conceptualization of the attack scenario, the threat can be quantified in the *Threat Matrix*. This comprises of the *Attack Potential, Motivation*, and *Impact*.

Attack Potential—a function of the system attack potential and the attacker's attack potential. System attack potential is the attack withstand capability of the system under attack. Attacker attack potential is attack actor's capability to conceptualize, launch and execute a successful attack. Variables for both components of attack potential incorporate: *Time Elapsed/Required* (time required for an attack actor for a successful attack and for the system withstand capability), *Finances* (how much the attack actor can access and how much is required to attack the system), *Expertise* (the level of expertise required by the attack actor), *Knowledge of the System* (the level of knowledge at the attacker's disposal and the knowledge required to attack the system), *Window of Opportunity* (the minimum access time required for the system under attack and maximum time available with the attack actor to execute the attack), and *Equipment* (the minimum tools required to scrutinize the system to execute a successful attack, the type of equipment required for the attack, and the type of equipment available to the attacker).

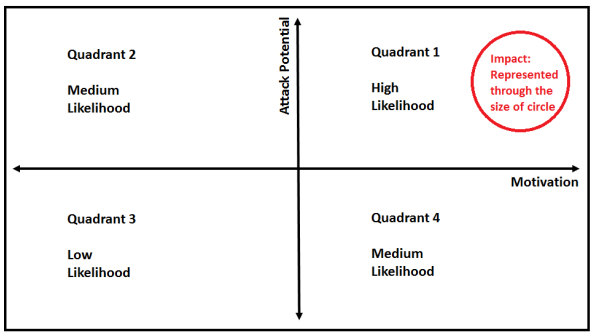

Fig. 2: Four quadrant plot.

Motivation—Financial Gain, Ideology, Passion, and Risk are considered as primary motivation factors. Financial gain describes financial benefits to the attacker after a successful cyber-attack on the automobile. This may include attacks of ransom, car theft, credit information theft from the infotainment system etc. Primary attack actors for this motivation are Thiefs and Organized Crime. Ideology, is a factor which motivates Hacktivists or Terrorists. A possible attack scenario would be hactivists attacking vehicles to protest against pollution caused by automobiles. Mechanics and Owners are primarily actors with Passion as the driving motivation. An attack scenario would be performance tuning by passionate owners and mechanics. The aforementioned motivation factors contribute positively towards cumulative motivation. Risk is a motivation factor which undermines the attack actors by providing the risk metric associated with a successful attack (e.g. jail term for a car thief).

Impact—the loss to the stakeholders. In the automotive case, this includes but is not not limited to Owners, Original Equipment Manufacturers (OEMs), Intelligent Transportation System (ITS) and Tier - I,II Suppliers, Auto Insurance Companies, and Government and Private entities involved in road infrastructure. Factors contributing to Impact are *Safety, Financial* and *Privacy*. Safety of stakeholders is violated when an attack hampers the safety applications and leads to devastating consequences and may include life or death scenarios. An attack scenario resulting in car theft, credit theft, or ransom-ware entail financial losses to stakeholders. Privacy loss includes identification of individual activities through location tracking or personal information gathering leading to blackmail.

Visualization—The tabular content of the Threat Matrix is presented through a visualization quadrant plot. An attack is represented as a circle centered at a point for which the horizontal axis quantifies the Attack Potential and the vertical axis quantifies the Motivation. Impact is represented through the size of the circle. The form of the quadrant plot is shown in

Fig. 2 and the respective quadrant representation is described in the figure.

IV. Use Case Analysis

With our reference architecture and threat model, we have a risk assessment framework with which we can assess an AD application. As an example, we explore the application of automated parking.

We begin by using our architecture to deconstruct the application. Automated parking is particularly representative because this is an application that can be realized at multiple levels of automation. At level 1, we have Parking Assist with Steering where a driver in the vehicle gets active steering assistance while parking. This application would use range sensors like ultrasonic proximity detectors and RADAR in addition to inertial and odometric sensors to perceive the vehicle's state relative to the environment. Using this information, the application would need to detect parked cars, free spaces, and the curb, localize the vehicle relative to other vehicles, and compute and execute a steering trajectory for a safe parking maneuver. At level 2, we have Key Parking as described in section II. Here, we introduce a key fob or smartphone with which the driver can initiate parking. This application would now involve computing and executing a combined steering, acceleration, and braking trajectory. At level 4, we have Driverless Valet Parking also as described in section II. For this level, the range sensor suite may be augmented with LIDAR or stereo cameras which are used with maps for precise localization and planning. Solutions may also consist of infrastructure like a central parking server to assign parking spots [12] or, in the case of Honda's proposed implementation, cameras for different views of the parking lot [13]. In all cases, we expect that the application restricts the driver from initiating parking until it senses that the current state is in its operational space.

From this framework, we can enumerate potential attacks and assess the risk associated with each attack. We note that there are many operational and implementation details which present entry points for attack such as maps [3] and the CAN bus. For our threat analysis, however, we shift our focus from the particulars of each attack surface to attacks which specifically exploit the application. Such attacks include:

Local Sensor Attacks

- *Blind range sensors* (low risk): An attacker could disable or blind the range sensors (either directly or remotely). We expect this to be of low risk because this attack is easily detected.
- *Message modification* (low risk): An attacker could alter sensor data in the system's internal network or provide false input to the range sensors to hallucinate objects or their lack thereof. We expect this attack to be of low risk as well because the first form of the attack would require physical access and the second, while feasible in a controlled setting [14], would require attacker expertise and be difficult to execute on a running system.

External Communication Attacks

- *Cloning remote control* (low risk): In the level 2 and level 4 scenarios, an attacker with access to the key fob or remote control could clone the device and be able to initiate the parking or un-parking maneuver at will. Even though the final effect of the attack can be performed remotely, this attack would still require physical access so we assign it low risk.
- *Spoof / replay parking signal* (high risk): An attacker could generate or replay a recorded parking signal from the key fob to achieve the same effect as above. We feel that this is a higher risk attack as all aspects of this attack could be performed remotely and the effect of this attack could be used toward car theft.
- *Spoof / replay valet signal* (high risk): Similar to the above attack, an attacker could compromise the level 4 application by spoofing the signal to retrieve a vehicle from its parking spot. If the application allowed for the vehicle to pick up the driver at their location, this kind of attack could be used to direct cars to an arbitrary location. While spoofing the valet signal could be challenging to execute, there is high motivation for a car thief to execute such an attack. Thus, we assign it high risk.

For the application of automated parking, car theft appears as a primary risk with several attacks exploiting the key and remote control. These attacks indicate that the vehicle must counter false external input by using internal sensors and protecting the key interface with strong multi-band authentication. Note that, in this example, the risks we associate with potential attacks and the conclusions we make are specific to our interpretations. The approach we propose is an adaptable framework and may result in different conclusions depending on the user's assumptions.

Although we demonstrated threat analysis with a single use case, more use cases will be presented. From this more comprehensive analysis, we note two interesting lessons. First, because remote access attacks are more likely, the wireless interfaces to the vehicle (e.g. DSRC, keys, map servers) are critical attack surfaces. As a result, AD systems must not necessarily trust information from the infrastructure or from other vehicles. Strong authentication and redundancy with local sensors are necessary to detect and combat potentially compromised wireless communications. Second, we learned from developing the reference architecture that, in terms of localization, fully automated systems tend to forgo GPS for highly precision and complex prior maps. GPS based attacks would likely have much lower impact on high level AD systems than attacks on maps or map servers.

V. Conclusion

We have proposed a risk assessment methodology for AD consisting of a generalizable reference architecture and a threat model to enable analysis of attacks that exploit automated driving applications. While we have demonstrated an example risk assessment and presented respective conclusions, the results of our framework are entirely customizable to a particular user's interpretations.

Even at this early stage, it is clear that AD applications will enable new and powerful attacks to safety and privacy. Opening up this discussion now is crucial for automakers and those invested in AD especially as we start to see partial and highly automated systems come to market. In future work, we aim to use our approach to develop a roadmap of cybersecurity risks to automated vehicles to understand how different threats become exposed as the technology evolves. We also plan to use our analysis to guide the design of prioritized security solutions to protect against high risk attacks.

REFERENCES

[1] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage, "Experimental Security Analysis of a Modern Automobile," in *2010 IEEE Symposium on Security and Privacy*. IEEE, 2010, pp. 447–462.

[2] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, "Comprehensive Experimental Analyses of Automotive Attack Surfaces," in *Proceedings of the 20th USENIX Conference on Security*. San Francisco, CA: USENIX Association, 2011, pp. 77–92.

[3] J. Petit and S. E. Shladover, "Potential Cyberattacks on Automated Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 1–11, 2014.

[4] C. McCarthy, K. Harnett, and A. Carter, "Characterization of Potential Security Threats in Modern Automobiles: A Composite Modeling Approach ((Report No. DOT HS 812 074)," National Highway Traffic Safety Administration, Washington, DC, Tech. Rep. October, 2014.

[5] O. Henniger, L. Apvrille, A. Fuchs, Y. Roudier, A. Ruddle, and B. Weyl, "Security requirements for automotive on-board networks," in *IEEE Transactions on Intelligent Transportation Systems*. IEEE, October 2009, pp. 641–646.

[6] SAE, "Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems," SAE International, Tech. Rep., 2014.

[7] A. Bartels, U. Eberle, and A. Knapp, "AdaptIVe Delivrable D2.1: System Classification and Glossary," Automated Driving Applications and Technologies for Intelligent Vehicles (AdaptIVe), Tech. Rep., 2015.

[8] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. N. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, M. Gittleman, S. Harbaugh, M. Hebert, T. M. Howard, S. Kolski, A. Kelly, M. Likhachev, M. McNaughton, N. Miller, K. Peterson, B. Pilnick, R. Rajkumar, P. Rybski, B. Salesky, Y.-W. Seo, S. Singh, J. Snider, A. Stentz, W. Whittaker, Z. Wolkowicki, J. Ziglar, H. Bae, T. Brown, D. Demitrish, B. Litkouhi, J. Nickolaou, V. Sadekar, W. Zhang, J. Struble, M. Taylor, M. Darms, and D. Ferguson, "Autonomous driving in urban environments: Boss and the Urban Challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, August 2008.

[9] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt, and S. Thrun, "Junior: The Stanford entry in the Urban Challenge," *Journal of Field Robotics*, vol. 25, no. 9, pp. 569–597, September 2008.

[10] A. Bacha, C. Bauman, R. Faruque, M. Fleming, C. Terwelp, C. Reinholtz, D. Hong, A. Wicks, T. Alberi, D. Anderson, S. Cacciola, P. Currier, A. Dalton, J. Farmer, J. Hurdus, S. Kimmel, P. King, A. Taylor, D. V. Covern, and M. Webster, "Odin: Team VictorTango's entry in the DARPA Urban Challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 467–492, August 2008.

[11] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, "Towards fully autonomous driving: Systems and algorithms," in *2011 IEEE Intelligent Vehicles Symposium (IV)*. Baden-Baden, Germany: IEEE, June 2011, pp. 163–168.

[12] P. Furgale, U. Schwesinger, M. Rufli, W. Derendarz, H. Grimmett, P. Muhlfellner, S. Wonneberger, J. Timpner, S. Rottmann, B. Li, B. Schmidt, T. N. Nguyen, E. Cardarelli, S. Cattani, S. Bruning, S. Horstmann, M. Stellmacher, H. Mielenz, K. Koser, M. Beermann, C. Hane, L. Heng, G. H. Lee, F. Fraundorfer, R. Iser, R. Triebel, I. Posner, P. Newman, L. Wolf, M. Pollefeys, S. Brosig, J. Effertz, C. Pradalier, and R. Siegwart, "Toward automated driving in cities using close-to-market sensors: An overview of the V-Charge Project," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, June 2013, pp. 809–816.

[13] Honda Motor Co., Ltd., "Overview of Honda Exhibit at the 20th ITS WORLD CONGRESS TOKYO 2013," 2013. [Online]. Available: http://world.honda.com/news/2013/4131008ITS-WORLD-CONGRESS-TOKYO/

[14] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl, "Remote Attacks on Automated Vehicles Sensors: Experiments on Camera and LiDAR," in *Black Hat Europe 2015*, Amsterdam, Netherlands, 2015.