CyberSecurity Considerations for an Interconnected Self-Driving Car System of Systems

Jeremy Straub, John McMillan, Brett Yaniero, Mitchell Schumacher,
Abdullah Almosalami, Kelvin Boatey, Jordan Hartman
Department of Computer Science
North Dakota State University
Fargo, ND, USA
jeremy.straub@ndsu.edu, john.mcmillan@ndsu.edu, brett.yaniero@ndsu.edu,
mitchell.schumacher@ndsu.edu, abdullah.almosalami@ndsu.edu, kelvin.boatey@ndsu.edu,
jordan.hartman@ndsu.edu

***Abstract*—The vehicle intercommunications required to enable the some of the most beneficial features of self-driving cars also pose significant security risks. To coordinate, cars must advise other cars of their plans, status and actions and be able to rely on the information provided by other cars. This paper presents an intrusion detection system, based on system-of-systems principles for the self-driving car system-of-systems.**

Keywords—self-driving cars; autonomous vehicles; car autonomy; cybersecurity; artificial intelligence; intrusion detection system; cyberphysical system security

I. Introduction

Self-driving cars promise significant benefits. They remove the driver from having to devote time and attention to commuting and may allow a passenger to sleep while his or her car carries him to a meeting in a different city and back. They also are posed to increase the efficiency and safety of the road system through vehicle to vehicle communications.

Cars, truck and other vehicles will be able to communicate their actions and plans to other nearby vehicles and transportation infrastructure systems. Vehicle to vehicle coordination may remove the need for stop signs and stoplights and the associated breaking and acceleration associated with them (and in the shorter term, reduce wait times). It also allows vehicles to alert other nearby vehicles to dangerous conditions and when the vehicle is taking an emergency response action. A vehicle can, for example, to others around it that it is breaking and those behind it can, nearly instantaneously, apply their own breaks as well to avoid an accident.

For this system to work, cars must send their details to other cars and be able to rely on the information provided by them. Problematically, each car is vulnerable to hacking by its owners or others and vehicle network failures may result in significant injury or death. It is, thus, desirable to develop systems to secure vehicles, communications networks and protocols used between them and detect prospective issues before they can cause incident or injury.

This paper proposes and presents an intrusion detection system for the self-driving car network system-of-systems. Techniques for detecting suspect activities and inputs are included. These range from detecting conventional attacks that are local to the vehicle all the way to detecting complicated information manipulation attacks designed to manipulate the actions taken by the vehicle. Through this detection, attacks can be automatically responded to and users can be made aware of the issue.

II. Background

Self-driving vehicle concepts range from limited-purpose self-parking [1] and operator attended driving systems [2], [3], which are currently available, to more robust autonomous and interconnected versions. Testing of more autonomous vehicles is ongoing [4]; however, the full power of self-driving vehicles won't be enjoyed until there are a multitude of autonomous and coordinating vehicles on the road.

These vehicles require path planning [5], motion estimation [6], position determination, robust sensing [7] and scheduling [8] capabilities. Even vehicles with human oversight or involvement (like systems that have a human operated lead car [9] or truck [10]) will require most of these features. Perhaps most importantly, they will need to be able to learn about their environment and user behavior – basically they need to be able to adapt everything from user routing preferences to decision making heuristics [11]. They can share this information and their plans with each other using vehicle ad-hoc networks (VANETs) [12] and benefit from doing so.

Even with the technical challenges that they pose, the economic value (from car sharing, reducing breaking and acceleration, reducing travel time and reducing the level of driver attention needed) that they are poised to provide is driving development. The benefits that they can provide to youth, the handicapped and the elderly [13] are also significant.

III. Self-Driving Car Intrusion Detection System Overview

The proposed intrusion detection system consists of four layers. The first is the vehicle-level intrusion detection. Each vehicle in the system of systems has its own basic IDS that allows it to function without access to the larger network. This IDS also performs analysis on information from the larger network to ensure that the trust placed in it is well-founded.

The second layer of IDS is the vehicle area network intrusion detection. This IDS works with VANETs which are comprised of several vehicles communicating and coordinating activities with one another. Each vehicle communicates with the others, transmitting data in real time and the IDS scans the data looking for anomalies and other signs of malicious activity.

The third level of the system handles wide area vehicle coordination and traffic management. This system secures information that goes beyond the local VANET. This information is collected from and transmitted to vehicles using road side units (RSUs). This wide-area system moves relevant information between VANETs and individual vehicles, as relevant. The IDS corresponding to this layer scans the information transmitted from each individual VANET to assess its reliability. Its goal is to determine if any VANETs, as a whole, have been compromised and if any nodes within the VANETs are compromised or malfunctioning. The IDS benefits from being able to compare, in some instances, data from one VANET to another (about the same vehicle, maneuver or phenomena) to look for manipulation and, generally, to ensure that the information being distributed is as accurate as possible.

The final layer of the IDS looks for more complex attacks that attempt to manipulate multiple areas or multiple layers of the self-driving vehicle control and coordination system. It is an overarching system that secures the interconnected network of RSUs and looks for critical issues with the system. This is a final failsafe that ensures that no widespread damage can occur from the malfunction of an RSU or a series of VANETs. This is done through a central processing node that correlates the data from the collection of RSUs.

IV. Car-Level Intrusion Detection

The first level of the IDS is the most basic in that it detects intrusions on the smallest scale, those that directed at the individual vehicle. This system is designed to operate without relying on network-based processing, so it doesn't seek to use comparative data or other remote resources and thus is limited in the attacks that it can detect and the symptoms that it considers. While it doesn't use networked resources, it can sense network traffic and seeks to evaluate whether the networked resources used for other system levels are trustworthy.

The system performs its analysis using data collected by the car. This data is used for both instantaneous analysis and to identify trends and patterns. For instance, the system may detect an attack meeting a known signature. Alternately, using the same data, it can collect data to characterize the typical day, activities and patterns of its users (schedule, routes, time at certain places, deviations from this and the reasons why, etc.). This data collection facilitates the creation of a baseline "normal" model. Commands that are deemed to be a significant deviation from the model would – potentially – indicate an intrusion and would be reviewed with greater scrutiny.

While not relying on a network to provide its core functionality, it does provide data to support other levels of intrusion detection. Abstractions of user normal models, which can help identify outlying behavior for users that don't use the system enough to have their own model, for example, could be contributed to the network-based intrusion detection systems. Additionally, the car IDS collects and supplies much of the car-level data analyzed by the network-based IDSs.

A. Mutually Verifiable Information

The concept of mutually verifiable information is a foundation of the car-level IDS. This process centers on information sent by other vehicles that can be verified by the IDS's vehicle's sensors. For instance, if the car ahead of the system's sends a signal that it is slowing down/breaking or speeding up, the user's car can verify this information (i.e., that this action is taken as stated) using its own sensors. Should the sent information not match the sensed data, the local sensor data is prioritized and the IDS will consider the sent information as suspect.

B. Verifying VANETs

The same principle is also applied to non-sensed information related to common protocols used between vehicles. One example is the VANET rating system for determining a car's reputation. In areas where a central system doesn't exist, a car would send its believed reputation along with historical data supporting it. The IDS would attempt to verify that the history matches the reputation purported and also verify elements of the history that it has data related to (i.e., operations in close proximity to the IDS's car). This would make reputation cheating (either as part of an attack or for other benefits that might be accorded to high reputation vehicles) a tedious endeavor because any piece of false information used could be potentially independently verified, severely damaging the vehicle's reputation.

C. IDS Verification

This same concept of reputation verification is applied to other intrusion detection systems as well. Using shared knowledge and information from multiple other systems (some of which have been verified with shared knowledge), defective or compromised IDSs can be discovered and trust in these systems can be adjusted appropriately.

D. Wrong Data and Compromise

It is critical to note that providing data which does not agree with directly sensed data of the local vehicle (or data that is believed true due to being provided by trusted sources) does not necessarily mean that a system is compromised or even malfunctioning. At the most benign, any sensing vehicle's vantage point may cause it to collect data that is not collected by another node. Differences in sensing capabilities, analysis methods and other factors may lead to sensed data being analyzed with different conclusions produced. Higher-level data products (which are based on analyzing conclusions based on sensed data may, similarly, differ based on the inputs provided to them). Because of this, the system does not make rapid trust/distrust decisions based on each data point evaluated. Instead, it gradually raises or lowers trust, based on the plurality of data points evaluated.

V. Vehicle Area Network Intrusion Detection

The VANET intrusion detection system is comprised of several levels of intrusion detection capabilities. The first utilizes the intrusion detection capabilities on the individual car level as previously described in Section V. The information from multiple individual cars is collected by the IDS on the local vehicle and analyzed. It is then transmitted to other nodes, which is the next level of the IDS. At this level, the system consists of several car-nodes that continuously communicate with each other. If a node's behavior is identified as suspicious, this finding is reported to the other nodes in the VANET which will then (typically, making their own decision to trust the report or not) avoid trusting the suspect node.

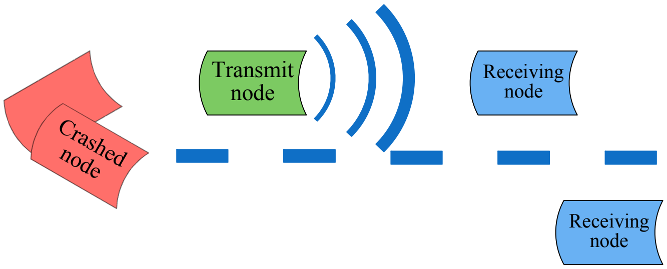

Figure 1. Example Scenario: Emergency Braking.

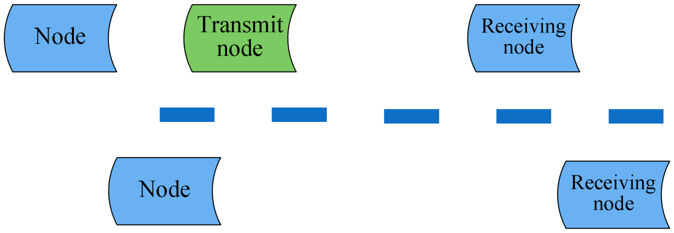

Figure 2. Example Scenario: False Emergency Braking.

At high levels of suspicion, a believed-malicious node will also be avoided on the roadway, in addition to having its data discounted. Figures 2 and 3 demonstrate the problem of false data by depicting one example of a malicious node trying to create an accident by alerting nodes to a fake one. Figure 1 shows what recipient nodes believe to be the case, from the report (where a leading vehicle detects an out-of-range roadway blocking accident). Figure 2 depicts the reality of the situation.

Within the VANET, each node is given a reputation rating, and its input is considered based on that rating. If the rating drops below a certain threshold, it is deemed suspicious. Falling below another (lower) threshold categorizes it as malicious.

As VANETs are created on an ad-hoc basis by nodes that are within a specific range of each other, problematic vehicles can be excluded. Deemed-suspicious nodes are isolated from the VANET and a warning message is broadcasted to other surrounding nodes. Deemed-Malicious nodes are retained in memory for exclusion longer and their warning is more broadly broadcast.

Alerts from suspicious and malicious nodes are not automatically routed, but serve as an advisory to perform verification activities to surrounding nodes. If the alert cannot be verified (or refuted), suspicious-node alerts are flagged as such and forwarded as advisories to other adjacent vehicle-nodes. Malicious-node alerts are dropped. Thus, in Figure 3, if node 6 were deemed to be malicious, nodes 1, 2 and 3 would be responsible for screening any alerts that it sent and any that could not be verified would be dropped. Note that, as many nodes may be members of multiple VANETs, non-dropped messages are widely spread (limited by a maximum relay count and other filtering.

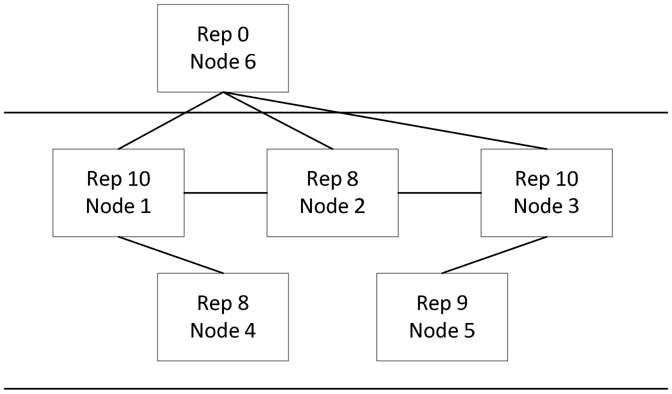

Figure 3. VANET Local Network Connections.

A. *Modular cross-layer intrusion detection system*

The proposed system is comprised of multiple modules. On each node, modules exist that are charged with collecting audit data on different IDS layers. With the aid of additional information from local hardware, a local decision-making module analyzes the audit data and determines when an activity is malicious or when it is normal.

Each node communicates with other nodes that are within a specific radius. It also indirectly communicates with nodes outside that radius but within the VANET. For example, in Figure 3, node 1 communicates with nodes 2, 4, and 6 because they are the closest nodes. Thus, if node 1 wants to send data to node 3, that message would be sent via either node 2 or 6. Deemed-suspicious and deemed-malicious nodes are not used for relaying, to the extent that other routes exist. Information relayed by suspicious nodes is treated with the same suspicion (triggering a validation attempt) as information generated by that node. Information relayed by deemed-malicious nodes is treated as an advisory and dropped if it cannot be confirmed.

The system's use of radius-limited messages prevents superfluous data transmissions. Node 1, for example, if sufficiently behind node 3, does not need to know what node 3 is doing (unless a critical-grade activity occurs or is detected).

Vehicles do not consider and will not relay information that they are beyond the radius-of-impact of. Some data is, of course, necessary for all nodes to know as quickly as possible, such as a major accident or malicious node being discovered. In this case, the node that discovers the issue chooses the most appropriate (safest, in the case of the malicious node – avoiding the node in route selection, as possible) route to broadcast the message.

B. Reputation

Each node is assigned a reputation rating from 1-10, based on the perception of its overall performance and reliability. The reputation is determined by the data sent by each node, and is decided by the surrounding nodes. A node cannot set or change its own reputation rating, it can only be affected by its output to other nodes. Reputation ratings are earned over time and are based on a window of performance.

New nodes start with no reputation and must develop reputation through their actions. However, as there is no pre-existing data, they can raise their score quickly. Thus, in Figure 3 and 4, node 6's reputation score of 0, could have a number of reasons:

- It is not a previously recognized node in the network
- It is broadcasting unreliable data

When nodes 1 and 2 read node 6's reputation rating, they will treat it as malicious and begin validating its data. Over time, if it is performing sufficiently, its reputation will rise and it will potentially become deemed-suspicious and, eventually, trusted.

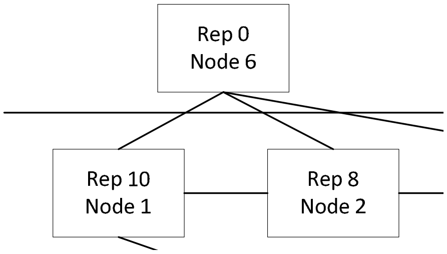

Figure 4. Zero reputation node.

Reputation ratings serve another important purpose, they allow nodes in the network to verify their data with well-established accuracy (those with the highest reputation rating in the network). Nodes with the highest reputations will typically be selected as the gateway nodes to other VANETs (proximity and IDS capabilities are also considered in this selection process as well). The nodes that are performing the best will aggregate the data from inside the VANET and send out the ad hoc network's performance statistics to other surrounding VANETs as well as to the central entity that aggregates and filters the data from all of the VANET's.

C. Local Decision Module

As shown in Figure 5, each car-node has a local decision module (LDM). Data from the local node as well as neighboring nodes are sent to the LDM of each node. The LDM is responsible for controlling the car and making all the decisions that the vehicle will execute. This system operates separately from, but not independently of the node's individual IDS. It correlates the data it receives on its own and processes alone and then juxtaposes it with the information received from the surrounding nodes. The LDM then, operates based on input received from neighboring nodes, whereas the individual IDS

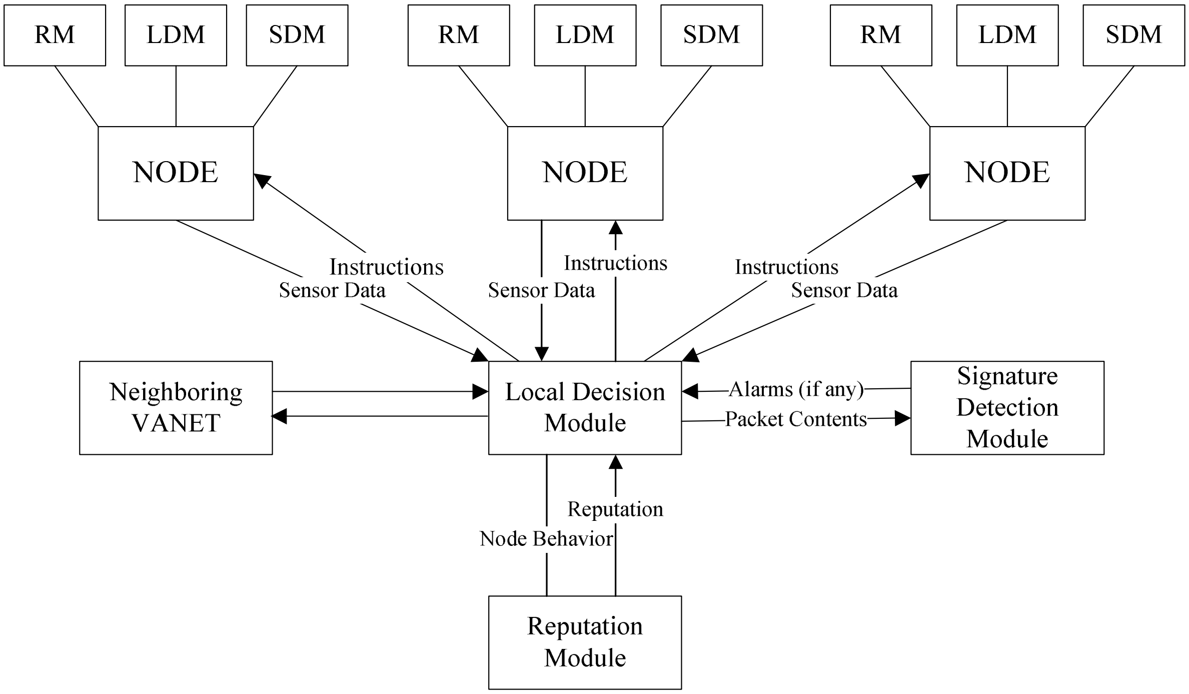

Figure 5. System overview.

will operate without input from other nodes. Based on that input, it also sends out data to other nodes that would be important for those nodes to receive.

D. Signature Detection Module

Received data is filtered through a system that is part of the IDS, known as the signature detection module (SDM). The SDM analyzes this data to identify the presence of any malicious signatures. If a known malicious signature is identified, then an alarm is sent to the LDM, which in turn sends that alarm to the appropriate nodes and to adjacent VANET's. If the SDM detects or receives any signature-matching malicious or anomalous data, the reputation of the node from which the data originated is impaired.

VI. Wide Area Vehicle Coordination & Traffic Management Intrusion Detection

The traffic management system within the IDS system of systems consists of two component systems, as depicted in Figure 6. The first is the road side units (RSUs). The second is its component on the individual vehicle within the VANETs,

A single RSU communicates with multiple VANETs, though the VANETs that an RSU coordinates with will, typically, change over time. Each RSU correlates the information it receives from nodes within the VANETs that it corresponds with. It filters out malicious node messages, based on malicious node identification from the VANETs and passes relevant information (including abstractions of information) to other VANETs over a wider area.

The RSU also looks for potentially malicious VANETs or nodes which may be mischaracterized as suspicious or malicious based on incorrect local information on the VANET. The RSU system also maintains longer-term history on nodes to facilitate analysis of malicious and suspicious node identification.

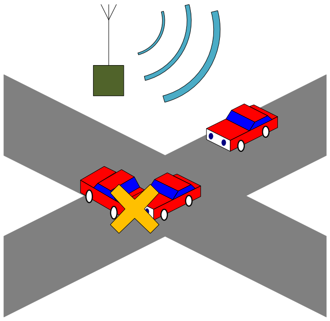

Figure 6. RSU communicating with local vehicles.

In addition to the data sharing for operational purposes, the RSUs also aids the IDS. The information sharing across RUSs and VANETs facilitate faster identification of malicious and suspicious nodes if they leave one VANET and enter another. Depending on proximity, the new VANET may already be (or could rapidly become) aware of the node's status and able to take appropriate action (as described previously). This would prevent or mitigate the potential misinformation and other damage it could do to a new VANET. Involving the RSU is also beneficial as it is able to perform a more thorough evaluation of the node, based on a larger dataset than a particular VANET could have access to, and also prevent VANETs that have been poisoned by malicious nodes from spreading the issue to other VANETs.

VII. Data Manipulation & Homeland Security

The vulnerabilities posed by self-driving cars extend beyond the risk to individual vehicles and their passengers. This section considers the potential that vehicles' interconnection and coordination mechanisms may be used to create a wide-scale attack. This type of attack would span VANETs and RSUs and might seek to compromise, pretend to replace, disable or confuse the large-scale coordination mechanism.

To combat these types of attacks, a large entity that is able to aggregate and process all of the data from these systems and communicate with the RSU's will need to be used in order to ensure that the system as a whole is not susceptible to widespread malicious or anomalous behavior. This capability is particularly important, as a successful attack at this level could have a large scale and turn vehicles into self-propelled weapons that injure both their occupants and those around them. Any large-scale attack could also cause a major disruption of regional or national travel capabilities.

The IDS to protect against this level of attack utilizes a combination of a centralized, hierarchical, and distributed mechanisms. It includes a central data handler that communicates with all RSUs (collecting information, in turn, from VANETs and individual nodes).

Previous sections have discussed how individual vehicles, VANETS and RSUs will validate upstream information, looking for an attack (or malfunction) that may provide inaccurate or malicious information. However, this detection is based on information that can be locally verified. A wider-scale attack, on the other hand, would seek to manipulate data beyond the ability of local validation. This could involve manipulating map data, GPS data or wide-scale configuration parameters or attempting to provide vehicles and RSUs with malicious software updates.

The central system, thus, seeks to check these data sources by looking for unexpected changes over time and anomalous and unexplained vehicle behavior patterns. It is inherently suspicious of both upstream (e.g., map, GPS) and downstream data providers (e.g., vehicles, RSUs).

While not inherently trusting the information provided, the wide-scale IDS benefits from local IDSs processing and detection capabilities. Any IDS can identify an issue that is beyond its local capability to investigate by providing a warning to the higher-level IDSs. Once an alert is generated, it is forwarded all the way to the centralized system. Intermediate IDSs process the warnings using their own (wider-scale) information and may conclude that an attack (or other malfunction) has or has not occurred. The IDSs may take local action based on a detected attack. They forward notifications of warnings (without taking immediate local action) to facilitate the identification of wider-scale attacks by higher-level IDSs. If higher-level IDSs detect a wide range attack (not identified by lower level IDSs), they will provide the lower-level IDSs with the information needed to confirm the attack. Most attacks at this scale would also trigger notification to system controllers to facilitate manual intervention and allow re-verification of system trust.

VIII. Qualitative Evaluation

The proposed system-of-system IDS for self-driving cars provides a logical foundation for future work. Each area of the system must, of course, be further elaborated on. Individual unit testing, system-to-system integration testing and testing of the entire system-of-systems will need to be performed.

One of the major considerations of the proposed IDS is that it relies upon the presence of RSUs for the levels beyond the individual vehicle. While VANETs could exist without RSUs, the ability to detect compromised VANETs and alert other VANETs about malicious and suspicious vehicles is impaired. Problematically, the feasibility of implementing RSUs cannot be adequately predicted at this time. Presuming that RSUs are a key component of self-driving vehicle systems (as opposed to some other approach being used), their deployment point in the self-driving vehicle system rollout is not yet determined. Also undetermined is who will pay for these units and how widely they will be deployed.

The range of the RSUs that will be deployed is also undetermined and could widely vary. This means that there is no way of predicting how many VANETs would typically be in communication with a single RSU at any given moment. In addition to this, there is no defined maximum membership for VANETs. The typical VANET size (and method of definition / enactment) will also have an impact on how much cross-VANET traffic will exist. The answers to these questions, when known, may require refinement of some assumptions of the system. However, they would not dramatically change it. The non-use of RSUs would require changes (potentially moving to a vehicle-to-central system direct or similar relationship).

The level of collaboration and coordination between vehicle manufacturers and their level of sharing regarding their self-driving architectures and algorithms will also impact how effective IDSs (which can better look, for example, for deception-based attacks when they know what information is relied upon) will be able to be in protecting the systems.

Finally, perhaps the largest unknown, is (beyond the car level), who will control the cybersecurity (including IDS) responsibility? Will it be national governments? State, regional and local governments? Vehicle manufacturers? A third party service provider? Cooperatives of vehicle owners? The chosen implementation may vary from area-to-area and this will undoubtedly impact the priorities of the cybersecurity (including IDS) system, its configuration and how it operates.

IX. Conclusions and Future Work

This paper has provided an overview of and detail related to several component systems that are part of a self-driving car intrusion detection system. A multi-homed, multi-level system with different system-of-systems component owners has been discussed and evaluated.

The goal of this paper has been to provide an overview of how these component systems can work together to provide a cohesive intrusion detection capability that augments other cybersecurity capabilities that will be incorporated. While it is expected that many of these other capabilities will form the first line of defense, the goal of the IDS is to act as a catch-all that prevents a prospective calamity that bypasses, compromises or simply evades the capabilities of other cybersecurity system elements. Plans or future work, thus, include continued effort in the development and testing of the IDS components and overarching system as well as testing in isolated and interconnected operational environments.

References

[1] M. Brown, "Tesla Autopilot's Latest Update Has Autonomous Parallel Parking | Inverse," *Inverse*, 24-Feb-2017.

[2] R. Duffer, "Tesla Model X P100D sets autopilot to a fast future - Portland Press Herald," *Portland Press Herald*, 17-Feb-2017.

[3] L. Mearian, "Here's why self-driving cars may never really be self-driving | Computerworld," *Computerworld*, 23-Feb-2017.

[4] T. Seppala, "Google's self-driving cars are getting better at autonomy," *Engadget*, 02-Feb-2017.

[5] U. Lee, S. Yoon, H. Shim, P. Vasseur, and C. Demonceaux, "Local path planning in a complex environment for self-driving car," in *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*, 2014, pp. 445–450.

[6] G. Hee Lee, F. Faundorfer, and M. Pollefeys, "Motion Estimation for Self-Driving Cars with a Generalized Camera," in *The IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 2746–2753.

[7] Gim Hee Lee, F. Fraundorfer, and M. Pollefeys, "Structureless pose-graph loop-closure with a multi-camera system on a self-driving car," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 564–571.

[8] J. Kim, H. Kim, K. Lakshmanan, and R. (Raj) Rajkumar, "Parallel scheduling for cyber-physical systems," in *Proceedings of the ACM/IEEE 4th International Conference on Cyber-Physical Systems - ICCPS '13*, 2013, p. 31.

[9] A. Wright and Alex, "Automotive autonomy," *Commun. ACM*, vol. 54, no. 7, p. 16, Jul. 2011.

[10] M. McFarland, "When truck drivers tailgating is actually a good thing - Feb. 16, 2017," *CNN Tech*, 2017. [Online]. Available: http://money.cnn.com/2017/02/16/technology/truck-platoons-peloton-omnitracs/. [Accessed: 26-Feb-2017].

[11] M. Bojarski *et al.*, "End to End Learning for Self-Driving Cars," Apr. 2016.

[12] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, "Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds," in *2014 IEEE World Forum on Internet of Things*, 2014.

[13] J. Yang and J. F. Coughlin, "In-vehicle technology for self-driving cars: Advantages and challenges for aging drivers," *Int. J. Automot. Technol.*, vol. 15, no. 2, pp. 333–340, Mar. 2014.

`