C2SMART
CONNECTED CITIES WITH SMART TRANSPORTATION
A USDOT University Transportation Center

New York University
Rutgers University
University of Washington
The University of Texas at El Paso
City College of New York

Developing Secure Strategies for Vehicular Ad hoc Networks in Connected and Autonomous Vehicles

May 2020

Developing Secure Strategies for Vehicular Ad hoc Networks in Connected and Autonomous Vehicles

Quanyan Zhu
Principal Investigator
New York University
0000-0002-0008-2953

Rae Zimmerman
Co-Principal Investigator
New York University
0000-0001-5825-3383

Song Fang
Postdoc
New York University

Siyu Zhou
Student
New York University

C2SMART Center is a USDOT Tier 1 University Transportation Center taking on some of today's most pressing urban mobility challenges. Some of the areas C2SMART focuses on include:

Urban Mobility and Connected Citizens

Urban Analytics for Smart Cities

Resilient, Smart, & Secure Infrastructure

Disruptive Technologies and their impacts on transportation systems. Our aim is to develop innovative solutions to accelerate technology transfer from the research phase to the real world.

Unconventional Big Data Applications from field tests and non-traditional sensing technologies for decision-makers to address a wide range of urban mobility problems with the best information available.

Impactful Engagement overcoming institutional barriers to innovation to hear and meet the needs of city and state stakeholders, including government agencies, policy makers, the private sector, non-profit organizations, and entrepreneurs.

Forward-thinking Training and Development dedicated to training the workforce of tomorrow to deal with new mobility problems in ways that are not covered in existing transportation curricula.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

The authors gratefully acknowledge support from the grant titled, "Developing Secure Strategies for Vehicular Ad hoc Networks in Connected and Autonomous Vehicles," from C2SMART, NYU Tandon School of Engineering through the U.S. Department of Transportation.
http://c2smart.engineering.nyu.edu/2019/06/14/secure-strategies-vehicular-ad-hoc-networks/

Executive Summary

A Vehicular Ad hoc Network (VANET) is a recent innovation in information and communication technologies to allow communications among vehicles and enable a connected and autonomous transportation system. Security and resilience are considered serious concerns.

The research aims at developing a resilient framework to be applied to transportation systems using connected and autonomous vehicles (CAVs). This innovation potentially responds to accident rates often related to inefficient communication systems, supported by a variety of state-of-the-art safety applications. A VANET is a self-organized, multi-purpose, service-oriented communication network enabling vehicle-to-vehicle and vehicle-to-roadside infrastructure communication for the purpose of exchanging messages to ensure an efficient and comfortable traffic system on roads. Its value, however, can potentially be impaired by cyberattacks. In particular, the focus of this research will be on false data injection attacks, in which a malicious agent aims at affecting CAV behavior by injecting in the network false information concerning, for example, the traffic condition in the area or the availability of charging stations for the CAVs.

Computational and analytical frameworks are developed to assess cyber risks of connected and autonomous vehicles (CAVs) and develop detection techniques based on learning algorithms. Countermeasures are then developed using anomaly identification techniques based on the learning and detection algorithms. In particular, the data collected from the nearby CAVs will be utilized to simulate and estimate the time evolution of the system; based on the outcome of such simulation, the CAV will apply a decision logic to distinguish between useful information and malicious data. In summary the approach aims at developing security game frameworks for vehicular networks that model the interaction between malicious attackers to VANETs and various defense mechanisms protecting them. Mitigation measures emphasizing multi-modal connectivity involving CAVs and/or traditional vehicles are integrated to promote resilience. This connectivity introduces flexibility in transport options by adapting and reconfiguring CAVs to existing multi-modal networks. Such flexibility also captures similar concepts for communication and information technologies such as redundant sensors and various kinds of backup systems and countermeasures.

This research is designed to provide guidance for users and developers of safe vehicular systems. The team's proposed work supports C2SMART's research goal in disruptive technologies, engagement and training and education. The project includes an outreach plan for industry, academia, professional and educational communities and an approach to technology transfer by planning to generate data sets and develop simulation tools and research results for the research community in a usable form and locating data sources in a data repository.

Table of Contents

Executive Summary iv
Table of Contents v
List of Figures vi
List of Tables vi
Section 1: Introduction 1
Section 2: Vulnerability and Risk Analysis and Consequence 4
Section 3: Cost-Effective Monitoring and Detection over VANET 7
Subsection 3.1 Introduction to Distributed 9
Subsection 3.2 Intrusion Detection and Monitoring over a VANET 10
Subsection 3.3 Distributed ML-based Detection over a VANET 11
Subsection 3.4 Numerical Evaluation 15
Section 4: Report on Resilient Multi-Modal Planning for Mitigation 19
Subsection 4.1 Frameworks for mitigation analysis 19
Subsection 4.2 Multi-modal Connections for Mitigation 20
Subsection 4.3 Summary 25
Section 5: Final Report and Publications 29
Section 6: Development of Outreach Plan 31
Subsection 6.1 Outreach objectives and approach 31
Subsection 6.2 Educational Outreach 31
Subsection 6.2 Industry Academic and Professional Community Outreach 32
Section 7: Technology Transfer, Data Repository, Guide and Demonstration Online 33

List of Figures

Figure 1. Trends in the Share of Cyber Attacks for Four Infrastructure Lifeline Sectors, 2012-2016 2
Figure 2: Simplified component security graph for an autonomous vehicle 5
Figure 3: Potential cyber risk graph of the autonomous vehicle components.....6
Figure 4: Individual risk scores and structural importance of components in the system. 6
Figure 5. A VANET scenario: TMC: traffic management center; V2V: vehicle-to-vehicle communications; V2I: vehicle-to- infrastructure communications; I2I: infrastructure-to-infrastructure communications. Each vehicle is equipped with an OBU and an AU.....8
Figure 6. Architecture of PML-CIDS: The Pre-processing engine collects and pre-processes the local audit data flow. The local detection engine then analyzes the pre-processed data using a classifier. If the user of the vehicle requires an update of the classifier, the P-CML engine is initiated. After collaborative learning, the updated classifier will be used in the intrusion detection.....11
Figure 7. Figure 7a-7b: convergence with different (fixed) values of av(t); DVP; non-private (Algorithm) 2) Figure 7c: security-privacy tradeoff measured by the false positive and false negative error rates; Figure 6d: Receiver operating characteristic (ROC) curve for non-private, DVP with different values of av.....16
Figure 8. Relationships among Existing Multi-Modal Arrangements and Potential Arrangements with CAVs by Trip Purpose 24

List of Tables

Table 1. Distribution of Trip Purpose Type Over Time, 2009 and 2017.....21
Table 2. Existing distribution of transportation mode by trip purpose, U.S.....22
Table 3. Hypothetical distribution of transportation mode by trip purpose assuming CAV substitution, U.S.....23

Section 1: Introduction

Road-traffic accidents claim thousands of lives every year in North America and all over the world, and may involve attributes of the transportation systems and the environments they are in. Statistics have shown that rural areas have typically dominated those mishaps, and in the last couple of years still account for a large share even though they have been declining [1: 4]. A Vehicular Ad hoc Network (VANET) is a significant innovation toward avoiding such deadly traffic mishaps with the assistance of a variety of state-of-the-art safety applications that integrate physical systems and user attributes. A VANET is a self-organized, multi-purpose, service oriented communication network enabling vehicle-to-vehicle (V2V) and vehicle-to-roadside infrastructure (V2I) communication for the purpose of exchanging messages to ensure an efficient and comfortable traffic system on roads ([2,3,4]. It is commonly anticipated that this network would play an effective role for active safety for roads and highways by introducing several different lifesaving applications for traffic management, driver safety, and drivers' assistance. Attention has been paid to portions of this issue, but greater emphasis is needed to refine the structure and function of the transportation – communication linkages with user or stakeholder input. The proposed research is a unique application and adapted design of VANET which incorporates communication and transportation linkages, cyber intrusions, and mitigation efforts involving multi-modal travel.

Numerous potential points for cyber interventions exist that have been identified in the literature for both connected vehicles and individual ICT-dependent vehicles and for roadways and roadway infrastructure. The number of such points in transportation that are ICT dependent is pervasive across many different functions such as switching, signaling, lighting, billing, tolling, and communications [5,6], and evidence exists that the deployment has been increasing [7,8,9]. Similarly, cyberattacks against transportation and communication infrastructure are considered significant among attacks on lifeline critical infrastructure sectors as shown in Figure 1. Transportation cyber-incidents increased as a percentage share of four infrastructure sectors from 2012-2015 with a decline in 2016 when communication began to increase dramatically [9: 14].

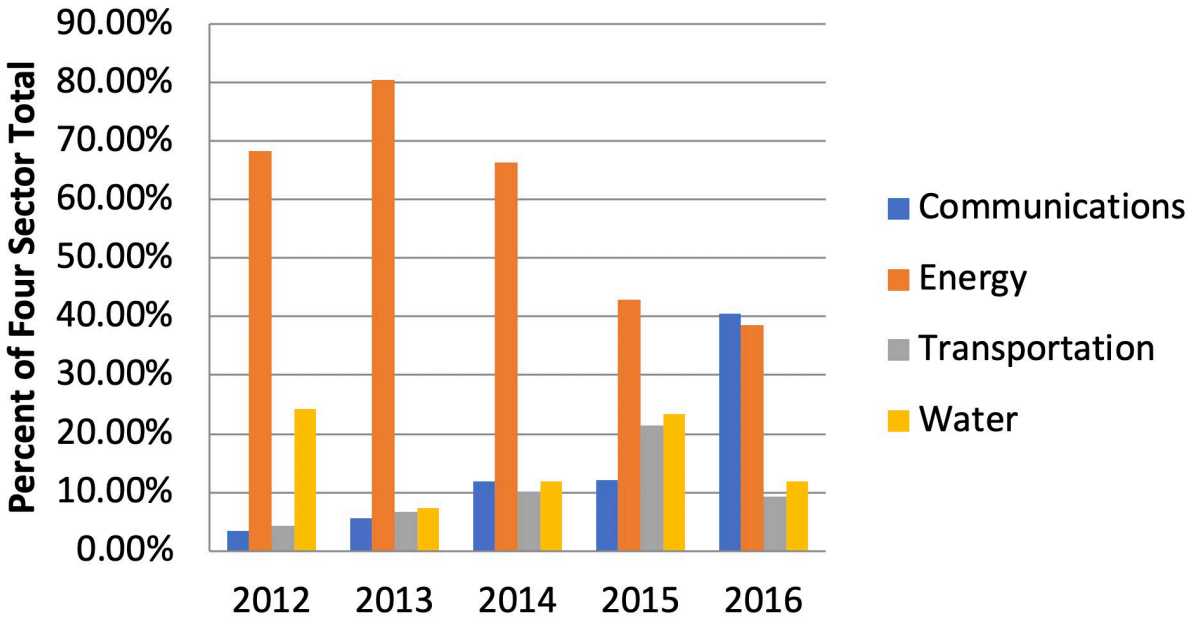


Figure 1. Trends in the Share of Cyber Attacks for Four Infrastructure Lifeline Sectors, 2012-2016

Source: Computed from the ICS-CERT databases from 2012-2016 by Zimmerman [9: 14]

The research presented here addresses vulnerable interdependence between transportation systems and the underlying critical infrastructure (emphasizing communication networks) supporting connected and autonomous vehicles (CAVs). The research focuses on intentionally malicious actions by humans whose intent is to impair or disrupt fundamental systems and services, in particular information systems, necessary to support CAVs. We assume CAVs make the V2V/V2I infrastructure a potential interdependent vulnerability between transportation and ICT.

The research aims at developing a resilient framework to be applied to transportation systems and the supporting ICT functions using connected vehicles. In particular, the focus will be on false data injection attacks, in which a malicious agent aims at affecting CAV behavior by injecting in the network false information concerning, for example, the traffic condition in the area or the availability of charging stations for the CAVs. The countermeasures will be developed using anomaly identification techniques based on learning and detection algorithms. In particular, the data collected from the nearby CAVs will be utilized to simulate and estimate the time evolution of the system; based on the outcome of such simulation, the CAV will apply a decision logic to distinguish between useful information and malicious data.

A critical part of the research is the integration of mitigation measures that reflect user and other stakeholder needs, namely, (1) multi-modal connectivity that improves network flexibility (2)

communication and information technology redundancies such as redundant sensors and (3) various kinds of backup systems and countermeasures. This research is designed to provide guidance for users and developers of safe vehicular systems.

Finally, as mentioned above, mitigation measures are designed and analyzed as to how they increase the resilience of the transportation systems against the impacts of cyber intrusions –through multi-modal connectivity involving both CAV and non-CAV vehicle technologies. As indicated earlier, the flexibility concepts multi-modal connectivity encompasses in the form of alternative routing and overrides is analogous to similar properties of other mitigation measures such as communications redundancy and backup systems and these extensions will be described.

References

[1] U.S. DOT National Highway Transportation Safety Administration "2017 Fatal Motor Vehicle Crashes: Overview", citing NHTSA's Fatality Analysis Reporting, 2018. System.https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812603

[2] Parno, B. and Perrig, A. Challenges in securing vehicular networks. In Workshop on hot topics in networks (HotNets-IV), 2005, pp. 1–6.

[3] Pathan, A.-S. K. Security of self-organizing networks: MANET, WSN, WMN, VANET, 2016. Boca Raton, FL, USA: CRC Press.

[4] Raya, M. and Hubaux, J.-P. Securing vehicular ad hoc networks. Journal of Computer Security, 2007, 15: 1, pp. 39–68.

[5] Zimmerman, R., Zhu, Q., Guo, Z. and Memon, N. "Overcoming Cyber Threats to Transportation Vehicles and Roadway Infrastructure," University Transportation Research Center's 3rd Annual Transportation Technology Symposium: Innovative Mobility Solutions, New York: NY Institute of Technology, 2015. http://files.ctctcdn.com/08b78404201/995ad396-a200-462c-8e1b-affa86866cf0.pdf

[6] Zimmerman, R. Transport, the Environment and Security. London, UK: Routledge, 2012.

[7] U.S. DOT, Status of the nation's bridges, highways, and transit conditions & performance, Washington, DC: U.S. DOT, 2008. From the ITS Deployment Statistics Database, Research and Innovative Technology Administration.

[8] U.S. DOT. Deployment of Intelligent Transportation Systems: A Summary Of the 2013 National Survey Results, 2014. FHWA-JPO-14-146. https://ntl.bts.gov/lib/54000/54200/54268/2013-National-ITS-Summary-Rpt_FINAL-7.pdf

[9] Zimmerman, R. “The cyber and critical infrastructures nexus: interdependencies, dependencies and their impacts on public services” White Paper, NYU Center for Cybersecurity, 2017. http://cyber.nyu.edu/wp-content/uploads/2017/12/Zimmerman-122717updateEdited-White-Paper-Cyber-and-Critical-Infrastructures.pdf

Section 2: Vulnerability and Risk Analysis and Consequence

Information and communications technology (ICT) systems are becoming increasingly complex, consisting of various different components connected together [1]. Often, these components are manufactured, controlled, or operated by different entities in different regions of the world. It is almost impossible to have centralized control over all entities in autonomous vehicles. Therefore, in addition to the conventional risks of system failures, there is another layer of risks emanating from different actors. With the proliferation of devices and networks, these risks are further amplified due to an unregulated and extremely heterogeneous ecosystem. Hence, it is becoming critical to develop methodologies to measure and analyze these risks. A more alarming concern is that the CAV systems directly interact with critical infrastructure systems leading to the possibility of cascaded failures and other disastrous consequences.

Risk analysis in CAVs is a challenge due to the lack of direct applicability of traditional methodologies in risk management. This challenge is primarily due to the structural complexity of the systems and the various different uncertainties that impact the assessment of systemic risk. Since fundamental design methodologies in technical industries rely heavily on modularity and abstraction, black-box systems that are difficult, if not impossible, to audit pose a constraint within which critical infrastructure security must be pursued. Processes to verify that components have been manufactured without hidden security vulnerabilities may be strategically important; however, the evident specialization required to produce complex ICT systems entails the need for a comparable specialization to provide credible verification. In general, then, acquirers or integrators of complex components must be prepared to assess and manage risks associated with trust. This challenging task calls for ongoing development of risk assessment methodologies and heuristics for evaluating the trustworthiness of devices [2,3]. When developing modeling practices suited to assessing these risks, it is necessary to consider the operational constraints within which model development proceeds. Given the uncertainties inevitably present in model development, processes should be designed to prioritize the clarification of uncertainties that pose the greatest threat to model utility. In this paper we present a brief graph-based model suited for the assessment of risks and develop case studies that illustrate the kinds of uncertainties that may impede the application of the modeling process to a given system. We then propose that structural uncertainties have the potential to be more significant than uncertainties in the probabilities of basic events. Similarly, among structural uncertainties, those pertaining to higher level components are particularly critical. Accordingly, the principle is developed that security risk assessment should prioritize

the accurate structural assessment of systems in order to improve the reliability of risk modeling in practice.

In this section, we aim to evaluate the risks of CAVs. The principal security risk we model involves attacks against the availability of safety critical components. The system graph in Figure 2 represents a hypothetical result of such a security analysis. Each minimal cut on the graph represents a complete attack. We assume here that a complete attack compromises the availability of any of the three terminal actuators. The supplier choice problem for this system involves choosing a supplier for each of the 15 components. For this case study we have chosen the following scenario for the supplier network. Each component may be acquired from one of three suppliers. No supplier offers more than one component. Of the 45 suppliers, there are five groups of five suppliers, with each group having a controller. The remaining 20 suppliers are independent. Although the supplier network topology described above is kept unmodified, parameters for risk, cost, supplier trust and group controller trust are varied across several different cases. In all cases, riskier components cost more. Generally, the cases range from simpler to more complex, with details of each described below along with results and discussion.

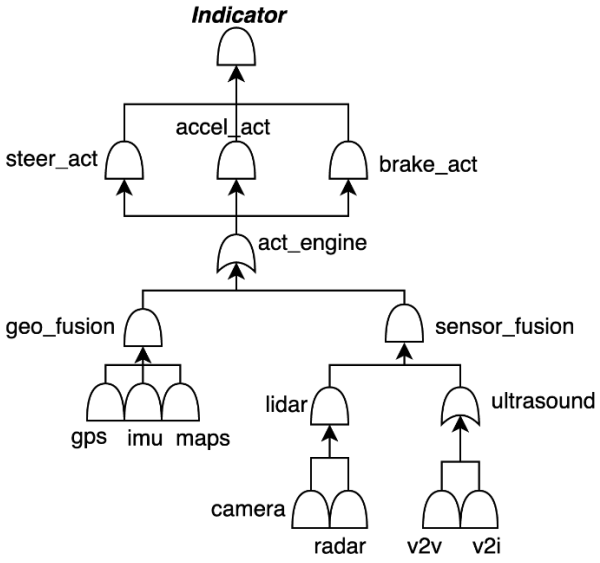


Figure 2: Simplified component security graph for an autonomous vehicle

To illustrate the use of the developed framework for risk analysis and mitigation decisions, we consider a case study of an autonomous vehicle system. The case study is based on a simplified model of the operation of an autonomous vehicle. Figure 2 shows a simplified representation of the logical interdependencies of various different components for the operation of an autonomous vehicle. The components include sensors such as radar, camera, GPS, IMU, etc. and actuators such as steering,

acceleration, and brake etc. System security is indicated by the top node 'Indicator'. A security failure of any of the actuators causes an incident. The principal security risk we model involves attacks against the availability of safety critical components. We assume here that a complete attack compromises the availability of any of the three terminal actuators.

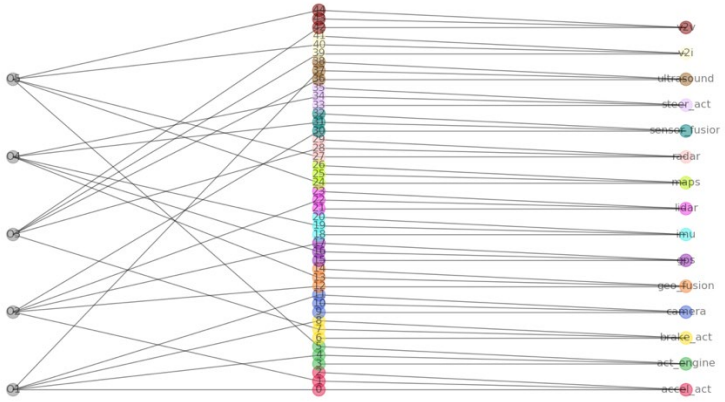
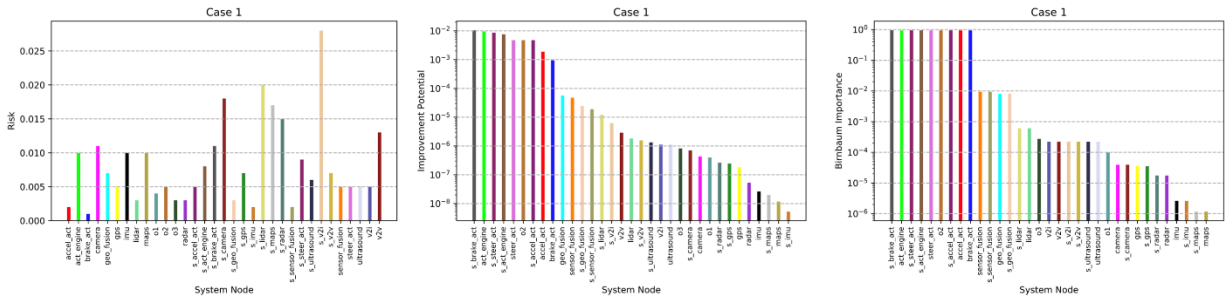


Figure 3: Potential cyber risk graph of the autonomous vehicle components

Figure 3 shows an example of the cyber risk graph of the components. Each component could be potentially sourced by three possible vendors. The vendors are in turn controlled by five main group owners. The risk propagates from the group owners via the vendors to individual components. The component risk then amalgamates through the system graph, shown in Figure 3, to form the systemic risk.



(a) Case 1 Risks (b) Case 1 Improvement Potential (c) Case 1 Birnbaum Importance

Figure 4: Individual risk scores and structural importance of components in the system.

In the developed case study, we have used a hypothetical risk profile of the suppliers, which results in a component-wise assessment of risk shown in Figure 4 (a). The overall systemic risk is a combination of these individual scores. An illustration of the criticality or role played by each component in the

operation of the autonomous vehicle is shown in Figure 4 (b) and (c). It is measured by the metrics of Improvement Potential and the Birnbaum Structural Importance measure. It can be observed that components with the highest risk may not have the most important structural importance in the system. Therefore, a holistic risk assessment is developed by taking these characteristics into account

This section was prepared by Quanyan Zhu [4].

References

[1] C. Folk, D. C. Hurley, W. K. Kaplow, and J. F. X. Payne, "The security implications of the Internet of things," AFCEA International Cyber Committee, Gaithersburg, MD, Tech. Rep., Feb. 2015.

[2] J. Boyens, C. Paulsen, R. Moorthy, and N. Bartol, "Supply chain risk management practices for federal information systems and organizations," National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep., Apr. 2015.

[3] C. Nissen, J. Gronager, R. Metzger, and H. Rishikof, "Deliver uncompromised: A strategy for supply chain security and resilience in response to the changing character of war," Mitre Corporation, Tech. Rep., Aug. 2018.

[4] Kieras, T., M. J. Farooq, and Q. Zhu under review "Modeling and Assessment of IoT Supply Chain Security Risks: The Role of Structural and Parametric Uncertainties," IEEE, under review. https://arxiv.org/pdf/2003.12363.pdf

Section 3: Cost-Effective Monitoring and Detection over VANET

With a growing number of vehicles on road and the rapid development of autonomous vehicles, road safety becomes an increasingly important issue. Vehicular ad hoc network (VANET) provides a communication system that enables the dissemination of safety-related information, traffic management, navigation, and road services. However, it is known that VANETs are vulnerable to a number of attacks from passive eavesdropping to active interfering [1]. For example, an attacker can eavesdrop and log the messages of other vehicles and replay them to access specific resources such as toll services. An attacker can intrude a specific vehicle, impersonate its identity, and send out false warnings that can disrupt the highway traffic [1].

Intrusion detection plays an important role in mitigating the threat of VANETs by using signature-based and/or anomaly-based approaches to detect adversarial behaviors [2]. Among many architectures of IDSs, the collaborative IDSs (CIDSs) have been proposed to enable the sharing of detection knowledge about known and unknown attacks and increase detection accuracy [3, 4, 5]. Distributed machine learning algorithms provide an appropriate framework for CIDSs to classify adversarial behaviors using

local datasets and share knowledge to increase the detection accuracy. In this paper, we consider the network-level intrusion attacks on computer system [6, 7] and take advantage of the collaborative nature of the VANETs and design a system architecture of a distributed machine-learning based CIDS over a VANET. The CIDS enables each vehicle to utilize the knowledge of the labeled training data of other vehicles; thus, it boosts the training data size for each vehicle without actually burdening the storage capacity of each vehicle.

Also, the laborious task of collecting labeled data can be distributed to all the vehicles in a VANET, thereby reducing the workload of each vehicle. Moreover, the CIDS enables the vehicles to share knowledge of each other without directly exchanging the training data. In addition, the CIDS provides the scalability of the training data processing and improves the quality of decision-making, while reducing the computational cost. The alternating direction method of multipliers (ADMM) [8] is one suitable approach to decentralizing the machine learning problem over a network that allows nodes over the network to share their classification results and yields the optimal classifier achieved under the centralized learning.

The main contributions of this work are the following. (i) We propose a machine-learning-based CIDS architecture to enable the collaborative information exchange and knowledge sharing in VANETs. (ii) We use ADMM to capture the distributed nature of a VANET and construct a collaborative learning over a VANET based on a regularized ERM algorithm.

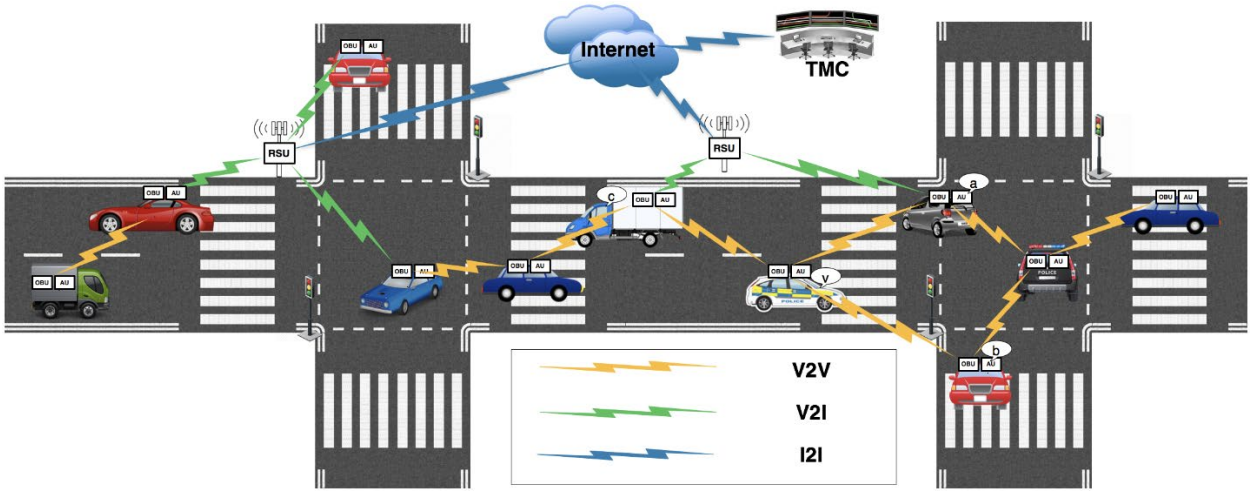


Figure 5. A VANET scenario: TMC: traffic management center; V2V: vehicle-to-vehicle communications; V2I: vehicle-to- infrastructure communications; I2I: infrastructure-to-infrastructure communications. Each vehicle is equipped with an OBU and an AU.

Subsection 3.1 Introduction to Distributed

Many works have studied various architectures of intrusion detection systems that are well-suited to MANET [3]. Most architectures for MANET can be classified into three categories. The first is the distributed and cooperative IDS, which captures the distributed nature of MANET that has the potential for constructing cooperation over the network. For example, Zhang and Lee in [10] have utilized this nature of MANET and constructed a model for a distributed and cooperative IDS. Also, Albers et al. have proposed a collaborative IDS based on local IDS by using mobile agents in [11]. The local IDS is implemented on each node of the MANET for local node-based security concerns, which can be extended to deal with the global security issues by establishing a collaboration among local IDSs over the MANET. The second category is hierarchical IDS model that extends the distributed and cooperative architectures. In [12], Sterne et al. have designed a dynamic hierarchical IDS using multilevel clustering. The third architecture uses the concept of mobile agents, which can move through the large network. In this type of framework, each mobile agent is assigned to work on a single specific task; then one or multiple mobile agents are distributed into each node in the MANET.

Previous research includes the work of Kachirski and Guha in [13] that has proposed distributed IDSs using multiple sensors based on mobile agent technology; and thus, the workload is distributed by separating functional tasks and assigning the tasks to different agents. Machine learning and data mining for IDSs have also been studied in the literature. These techniques enable the IDS to continuously learn attacks and their behaviors, enhance the knowledge of the security system, make connections between suspicious events, and predict the occurrence of an attack. Researchers have studied the unsupervised learning such as the technique of clustering, which is an unsupervised pattern discovery method, in IDSs. There are several approaches for clustering the unlabeled data; for example, Blowers and Williams [14] have applied a density-based spatial clustering of applications with noise clustering algorithm to group normal versus anomalous network packets.

Other clustering-based work includes hierarchical clustering [15] and K-means [16]. There is also literature on the IDS with supervised learning such as support vector machine [17]. For example, Wagner et al. [18] have applied one-class SVM classifier and used a new window kernel to find the anomaly based on time position of the data. Other methods using supervised learning include decision trees [19, 20, 21], artificial neural networks [22, 23], and sequential data aggregation [24, 25]. There also have been works on intrusion-prediction based detection using non-machine-learning techniques [26, 27]. For example, Nidhal et al. [28] have designed a game-theoretic intrusion detection approach for VANET. The game-based model can predict a possible future denial-of-service attack on the monitored nodes.

Algorithm 1 PML-CIDS

Input: Real-time VANET system data: Local audit data flow and activity logs.

Step 1: The ***pre-processing engine*** collects and pre-processes the real-time VANET system data, by numerical transformation, features selection, and data normalization.

if the classifier needs update **then**

Step 2: The ***P-CML engine*** is initiated and local training dataset is loaded. And updated classifier is obtained.

Step 3: The ***local detection engine*** uses the newly updated classifier to analyze the real-time VANET system data. If any activities are classified as intrusions, the ***local detection engine*** triggers the alarm.

Step 2: The ***local detection engine*** uses the current classifier to analyze the real-time VANET system data and triggers the alarm when any activities are classified as intrusions.

end if

Subsection 3.2 Intrusion Detection and Monitoring over a VANET

In this section, we describe the architecture of the proposed PML-CIDS which includes multiple building blocks for VANETs. Illustrated in Figure 1, a general VANET consists of on-board units (OBU), application units (AU), and roadside units (RSU). The communication between OBUs (vehicle-to-vehicle), or between an OBU and an RSU (vehicle-to-infrastructure) is based on wireless access in-vehicle environment (WAVE) [3]. The RSUs can also connect to other infrastructures such as other RSUs and traffic management center, and the communications between them (infrastructure-to-infrastructure) are through other wireless technology. Each vehicle is equipped with an OBU and one or multiple AUs. It also has a set of sensors to collect information and use the OBU to exchange information with other OBUs or RSUs. Details about the three main components of the VANET architecture are presented in the Appendix A.1 for interested readers.

Each vehicle is equipped with one local PML-CIDS agent as shown in Figure 2 to monitor its local activities including the ones in the AU and the communications via the OBU. Conceptually, the collaborative system consists of three main components, namely, pre-processing engine, a local detection engine, and privacy-preserving collaborative machine learning (P-CML) engine. The logical flow of a PML-CIDS is illustrated in Algorithm 1. The pre-processing engine gathers and pre-processes the real-time VANET system data that describe the system activities in a vehicle. The pre-processed system data is then analyzed by the local detection engine using classification techniques. If the user of the vehicle requires the current classifier to be updated, then the P-CML engine is initiated. The local

detection engine uses the newly retained classifier to analyze the system data. Otherwise, the current classifier is used in the classification of intrusions. If any intrusion is classified, the alarm is triggered. One essential component of the CIDS is the P-CML engine which is composed of the collaborative communication (CC) engine, distributed local learning (DLL), and privacy-preserving (PP) mechanism.

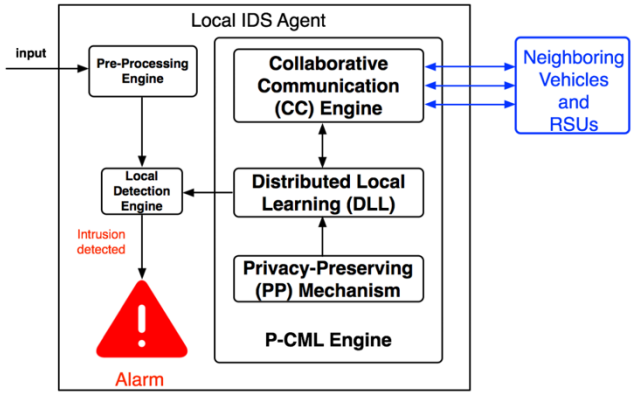


Figure 6. Architecture of PML-CIDS: The Pre-processing engine collects and pre-processes the local audit data flow. The local detection engine then analyzes the pre-processed data using a classifier. If the user of the vehicle requires an update of the classifier, the P-CML engine is initiated. After collaborative learning, the updated classifier will be used in the intrusion detection.

Subsection 3.3 Distributed ML-based Detection over a VANET

Consider a connected VANET, which consists of P vehicles, described by an undirected graph G(V , E) as shown in Figure 3 at time t, with the set of vehicles V = {1,2,3,...,P}, and a set of edges E denoting the links between connected vehicles. In general, the graph can change over time as the nodes move. Here we introduce the framework with a fixed topology, and it can be easily extended to dynamic regimes. A particular vehicle v ∈ V only exchanges information between its neighboring vehicle w ∈ Nv, where Nv is the set of all neighboring vehicles of v. Each vehicle stores a labeled training dataset Dv = {(xiv,yiv) ⊂ X ×Y : i = 0,1,⋯ ,nv} of size nv, where xiv ∈ X ⊆ Rd and yiv ∈ Y := {−1, 1} are the data vector and the corresponding label, respectively. The entire network therefore has a set D. The training dataset Dv of v ∈ V contains data points describing the VANET system activities such as user and application activities and communication activities through the OBUs; each data point is labeled as an intrusion (1) or a normal activity (−1). Labeling is important, which is even true for the pure anomaly detection [17]. The

labeled training dataset must have both the normal data and the intrusion data including the novel attack data.

The collaborative learning in our model should be distributed over a VANET without direct data sharing. The alternating direction method of multipliers (ADMM) is a suitable approach for our model. In this work, we focus on a class of distributed ADMM-based empirical risk minimization (ERM) as the supervised learning algorithm used in the collaborative learning. Each collaborative learning can be modeled as an optimization problem to find a classifier $f : X \rightarrow Y$ using all available data D that enable all vehicles in the ad hoc network to classify any input data x′ (i.e., data flow collected and pre-processed by pre-processing engine) to a label $y' \in \{-1, 1\}$, where −1 and 1 denote normal activities and intrusions, respectively. Before introducing the ADMM-based distributed learning, we first present the centralized optimization in the next subsection.

Let Z1 (f |D) be the centralized objective function of a regularized empirical risk minimization problem (C-ERM). Thus, the C-ERM problem can be defined as:

$$\text{Min } Z_1 := C_1/n_v \sum_v \sum_i L(y_{iv}, f_v x_{iv}) + P \sum_v k R(f_v) \quad (1)$$

where C1 ≤ nv is a regularization parameter and κ > 0 is the parameter that controls the impact of the regularizer. The regularizer function R(f) in (1) is used to prevent overfitting. Suppose that D is available to the fusion center vehicle, a global classifier $f : X \rightarrow Y$ is chosen by optimizing the C-ERM.

To solve the problem by ADMM, we first decentralize the C-ERM problem by introducing the decision variables { fv}Pv=1; then, vehicle v determines its own classifier fv. We impose consensus constraints f1 = f2 = ... = fP to guarantee the global consistency of the classifiers. Let {svw} be the auxiliary variables to decouple fv of the vehicle v from its neighbors $w \in N_v$ in the VANET. Then, the consensus-based reformulation of C-ERM becomes

$$\text{Min } C_1/n_v \sum_v \sum_i L(y_{iv}, f_v x_{iv}) + P \sum_v k R(f_v)$$

$$\text{s.t. } f_v = s_{vw},\ s_{vw} = f_v,\ v = 1, \ldots, P,\ w \in N_v \quad (2)$$

where Z2({ fv}v∈V |D) is the reformulated objective as a function of { fv}v=1. Problem (2) can be solved in a distributed fashion using ADMM with each vehicle v ∈ V optimizing the following distributed regularized empirical risk minimization problem (D-ERM): Zv:=C1/nv ∑i L(yiv , fv xiv) + p R(fv)

(3)

The augmented Lagrange function associated with the D-ERM is:

$$L_v(f_v, s_{vw}, \lambda_{vw}) = Z_v + \lambda_{vi}^{a} (f_v - s_{vi}) + \sum_i (\lambda v_i^{b})^T (s_{vi} - f_i) + \eta/2 \sum_i (|f_v - s_{vi}|^2 + |s_{vi} - f_i|^2)$$

(4)

Therefore, the distributed iterative procedures to solve (3) are:

$$f_v(t + 1) = \text{argmin } L_v^D(f_v, s_{vw}(t), \lambda_{vw}^k(t)),$$

(5)

$$s_{vw}(t + 1) = \text{argmin } L_v^D(f_v(t + 1), s_{vw}, \lambda_{vw}^k(t)),$$

(6)

$$\lambda_{vw}^a(t + 1) = \lambda_{vw}^a(t) + \eta(f_v(t + 1) - s_{vw}(t + 1)), v \in V, w \in N_v,$$

(7)

$$\lambda_{vw}^b(t + 1) = \lambda_{vw}^b(t) + \eta(f_v(t + 1) - s_{vw}(t + 1)), v \in V, w \in N_v,$$

(8)

Here, svw (t + 1) in (6) can be found in closed form because the cost in (6) is linear-quadratic in svw (t + 1)[39]. By substituting the closed-form solution, we can eliminate svw (t + 1) from LvD ; this approach makes it possible to simplify the iterative procedures (5) to (8). Indeed, according to Lemma 3 in [39], we can further simplify the distributed iterative procedures by initializing the dual variables λ k = 0 and combining the two sets of dual variables into one as λv(t) = ∑w∈Nv λvkw, v ∈ V , w ∈ Nv, k = a, b. Then, we can combine (7) and (8) into one update. We simplify (5)-(8) by introducing the following. Let LvN(t) be the short-hand notation of LvN({fv},{fv(t)},{λv(t)}) as:

$$L_v^N := C_1/n_v \sum_i L(y_{iv}, f_v x_{iv}) + p\, R(f_v) + 2\lambda_v(t)^T f_v + \eta \sum_i |f_v - 1/2(f_v(t) - f_i(t))|^2,$$

(9)

The ADMM iterative procedures (5)-(8) are reduced to

$$f_v(t + 1) = \text{argmin } L^N(f_v, f_v(t), \lambda_v(t)),$$

(10)

$$\lambda_v(t + 1) = \lambda_v(t) + \eta/2 \sum_w [f_v(t + 1) - f_w(t + 1)].$$

(11)

Algorithm 2 Distributed ERM over VANET

Required: Randomly initialize f_v, $\lambda_v = \mathbf{0}$

Input: $\hat{D}$

for $t = 0, 1, 2, 3, \ldots$ **do for** $v = 0$ **to** P **do**

Compute $f_v(t+1)$ via (10).

for $v = 0$ **to** P **do**

Broadcast $f_v(t+1)$ to all neighboring vehicles $w \in N_v$.

for $p = 0$ **to** P **do**

Compute $\lambda_v(t+1)$ via (11).

end for

Output: $f^* = f_v$, for all $v \in V$.

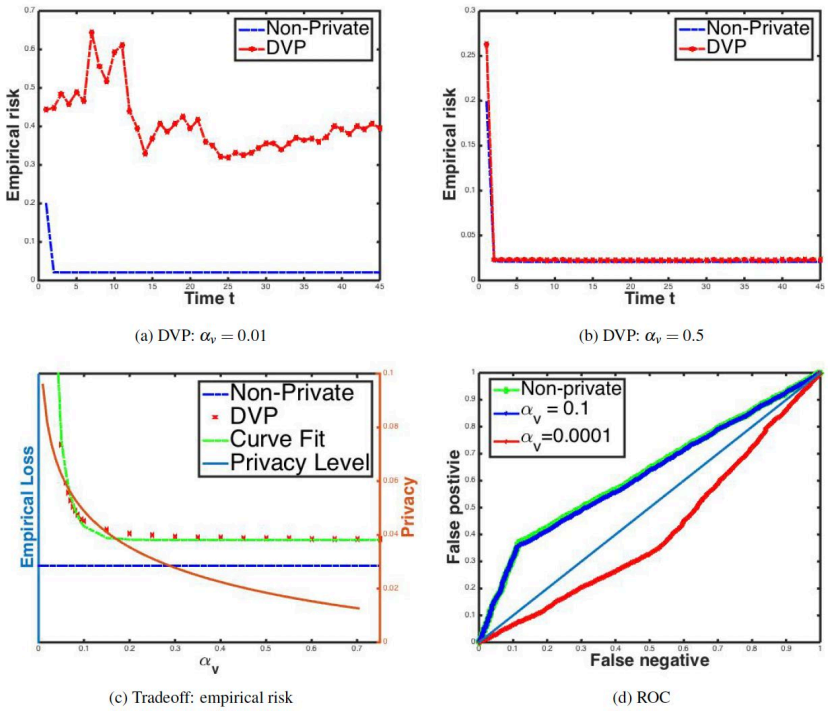
Algorithm 2 summarizes the (non-private) distributed ERM over a VANET. At iteration t + 1, vehicle v updates its local fv(t) through (10). Next, v broadcasts the latest fv(t + 1) to all its neighboring vehicles w ∈ Nv. When each vehicle has updated λv(t+1) via (11), iteration t+1 finishes. Throughout the entire algorithm, each vehicle v∈V only updates its own fv(t) and λv(t) and the only information exchanged between neighboring vehicles is fv(t); thus, direct data sharing is avoided. There are several methods to solve (10). For example, projected gradient method, Newton method, and Broyden-Fletcher-Goldfarb-Shanno method [40] that approximates the Newton method, to name a few. In this distributed algorithm, each vehicle solves a minimization problem per iteration using its local training dataset. The only information in the message transmitted by the OBUs between neighboring vehicles is the value of fv(t).

ADMM-based distributed machine learning has benefits due to its high scalability. It also provides a certain degree of privacy since vehicles do not share training data directly. However, the privacy issue arises when powerful adversaries can make intelligent inferences at each step of the collaborative learning and extract the privacy information contained in the training dataset based on their observation of the learning output of each vehicle. Simple anonymization or conventional sanitization is not sufficient to address the privacy issue as mentioned in the introduction. In the next subsection, we will discuss the privacy concerns about the training data and propose differential privacy solutions.

Subsection 3.4 Numerical Evaluation

In this section, we test the learning performance of Algorithm 3 and explore the tradeoff between security and privacy. We simulate the user and system activities, the communication activities in the AUs and OBUs of the VANET based on the NSL-KDD data, which is the refined version of its predecessor of KDD'99 and solves some of the inherent problems of the KDD'99 [44]. The NSL-KDD dataset contains essential records of the complete KDD dataset. Each record contains 41 attributes indicating different features of flow with a label assigned either as an attack or normal. Due to the lack of public datasets for network-based IDSs, the NSL-KDD is currently the best available dataset for benchmarking of different intrusion detection methods [17, 29]. In the experiments, the task is to classifier whether a network activity is an attack (1) or normal (-1) using logistic regression. There are four types of attacks presented in NSL-KDD, namely, denial of service, probing, unauthorized access to local system administrator privileges, and unauthorized access from a remote machine [30]. In this experiment, we only classify whether an activity is an attack or normal without identifying the specific type of the attack. We also propose an approach to select an optimal value of av(t) that can manage the tradeoff between security and privacy by introducing a utility function of privacy. In the experiments, we fix the value of av(t) for each entire running of Algorithm 3; thus, the noise of each vehicle generated at each running of DVP is i.i.d. To process the NSL-KDD dataset into a form suitable for the classification learning, we process the NSL-KDD dataset according to the procedures suggested in [31]. The main processes include the

transformation of symbolic attributes to numeric values, feature selection that eliminates irrelevant, noisy or redundant features, data normalization that helps speed up the learning. When the P-CML engine in vehicle v is initiated, collaboration is established over the VANET. As shown in Figure 5, the vehicle only communicates with vehicles in its one-hop neighborhood composed of three vehicles, a, b, and c, which also communicate with neighboring vehicles directly or through the RSU (e.g., a and c). Each vehicle in the collaboration updates its own primal and dual parameters simultaneously.



(a) DVP: $\alpha_v = 0.01$ (b) DVP: $\alpha_v = 0.5$

(c) Tradeoff: empirical risk (d) ROC

Figure 7. Figure 7a-7b: convergence with different (fixed) values of av(t); DVP; non-private (Algorithm) 2) Figure 7c: security-privacy tradeoff measured by the false positive and false negative error rates; Figure 6d: Receiver operating characteristic (ROC) curve for non-private, DVP with different values of av.

This section was prepared by Quanyan Zhu [32].

References

[1] A.-S. K. Pathan, Security of self-organizing networks: MANET, WSN, WMN, VANET. CRC press, 2016.

[2] W. Zhang, R. Rao, G. Cao, and G. Kesidis, "Secure routing in ad hoc networks and a related intrusion detection problem," in Military Communications Conference, 2003. MILCOM'03. 2003 IEEE (2003), pp. 735–740, 2003.

[3] T. Anantvalee and J. Wu, "A survey on intrusion detection in mobile ad hoc networks," in Wireless Network Security, pp. 159–180, Springer, 2007.

[4] Q. Zhu, C. Fung, R. Boutaba, and T. Basar, "Guidex: A game-theoretic incentive-based mechanism for intrusion detection networks," IEEE Journal on Selected Areas in Communications, 30: 11 (2012), pp. 2220–2230.

[5] C. J. Fung, Q. Zhu, R. Boutaba, and T. Basar, "Bayesian decision aggregation in collaborative intrusion detection networks," in Network Operations and Management Symposium (NOMS), IEEE (2010), pp. 349–356, IEEE, 2010.

[6] J. Raiyn et al., "A survey of cyber-attack detection strategies," International Journal of Security and Its Applications, 8:1 (2014), pp. 247–256.

[7] N. Hoque, M. H. Bhuyan, R. C. Baishya, D. K. Bhattacharyya, and J. K. Kalita, "Network attacks: Taxonomy, tools and systems," Journal of Network and Computer Applications, 40 (2014), pp. 307–324.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," Foundations and Trends in Machine Learning, 3:1 (2011), pp. 1–122.

[9] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in Theory of cryptography, pp. 265–284, Springer, 2006.

[10] Y. Zhang, W. Lee, and Y.-A. Huang, "Intrusion detection techniques for mobile wireless networks," Wireless Networks, vol. 9, no. 5, pp. 545–556, 2003.

[11] P. Albers, O. Camp, J.-M. Percher, B. Jouga, L. Me, and R. S. Puttini, "Security in ad hoc networks: a general intrusion detection architecture enhancing trust based approaches.," in Wireless Information Systems (2002), pp. 1–12.

[12] D. Sterne, P. Balasubramanyam, D. Carman, B. Wilson, R. Talpade, C. Ko, R. Balupari, C.-Y. Tseng, and T. Bowen, "A general cooperative intrusion detection architecture for VANETS," in Third IEEE International Workshop on Information Assurance (IWIA'05), IEEE (2005), pp. 57–70.

[13] O. Kachirski and R. Guha, "Effective intrusion detection using multiple sensors in wireless ad hoc networks," in System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference, IEEE (2003), pp. 8.

[14] M. Blowers and J. Williams, "Machine learning applied to cyber operations," in Network Science and Cybersecurity, pp. 155-175, Springer, 2014.

[15] S.-J. Horng, M.-Y. Su, Y.-H. Chen, T.-W. Kao, R.-J. Chen, J.-L. Lai, and C. D. Perkasa, "A novel intrusion detection system based on hierarchical clustering and support vector machines," Expert systems with Applications, 38:1 (2011), pp. 306–313.

[16] Z. Muda,W. Yassin, M. Sulaiman, and N. Udzir, “Intrusion detection based on k-means clustering and naïve bayes classification,” in Information Technology in Asia (CITA 11), 2011 7th International Conference, IEEE (2011), pp. 1–6.

[17] A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Communications Surveys & Tutorials, 18:2 (2015), pp. 1153–1176.

[18] C. Wagner, J. François, T. Engel, et al., “Machine learning approach for ip-flow record anomaly detection,” in International Conference on Research in Networking, pp. 28–39, Springer, 2011.

[19] C. Kruegel and T. Toth, “Using decision trees to improve signature-based intrusion detection,” in International Workshop on Recent Advances in Intrusion Detection, pp. 173–191, Springer, 2003.

[20] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi, “Exposure: Finding malicious domains using passive dns analysis.,” in NDSS, 2011.

[21] L. Bilge, S. Sen, D. Balzarotti, E. Kirda, and C. Kruegel, “Exposure: a passive DNS analysis service to detect and report malicious domains,” ACM Transactions on Information and System Security (TISSEC), 16:4 (2014), p. 14.

[22] J. Cannady, “Artificial neural networks for misuse detection,” in National information systems security conference (1998), pp. 368–81.

[23] R. P. Lippmann and R. K. Cunningham, “Improving intrusion detection performance using keyword selection and neural networks,” Computer Networks, 34: 4 (2000),pp. 597–603, 2000.

[24] C. J. Fung and Q. Zhu, “FACID: A trust-based collaborative decision framework for intrusion detection networks,” Ad Hoc Networks, 53 (2016), pp. 17–31, 2016.

[25] Q. Zhu, C. J. Fung, R. Boutaba, and T. Basar, “A distributed sequential algorithm for collaborative intrusion detection networks,” in Communications (ICC), 2010 IEEE International Conference, IEEE (2010), pp. 1–6,

[26] Q. Zhu and T. Basar, “Dynamic policy-based ids configuration,” in Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference, IEEE (2009), pp. 8600–8605.

[27] Q. Zhu and T. Basar, “Indices of power in optimal ids default configuration: Theory and examples,” in International Conference on Decision and Game Theory for Security, Springer (2011), pp. 7–21.

[28] M. N. Mejri, N. Achir, and M. Hamdi, "A new security games based reaction algorithm against dos attacks in vanets," in Consumer Communications & Networking Conference (CCNC), 2016 13th IEEE Annual, IEEE (2016), pp. 837–840.

[29] M. Tavallaee, E. Bagheri, W. Lu, and A.-A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defense Applications 2009 (2009).

[30] L. Dhanabal and D. S. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms," International Journal of Advanced Research in Computer and Communication Engineering, 4:6 (2015).

[31] H. Mohamad Tahir, W. Hasan, A. Md Said, N. H. Zakaria, N. Katuk, N. F. Kabir, M. H. Omar, O. Ghazali, and N. I. Yahya, "Hybrid machine learning technique for intrusion detection system," 5th International Conference on Computing and Informatics (ICOCI) 2015 (2015).

[32] T. Zang and Q. Zhu "Differentially private collaborative intrusion detection systems for VANETS," IEEE, under review. https://arxiv.org/pdf/2005.00703.pdf

Section 4: Report on Resilient Multi-Modal Planning for Mitigation

Subsection 4.1 Frameworks for mitigation analysis

The overall project goals for mitigation include the development of a resilient framework to be applied to transportation systems using connected and autonomous vehicles (CAVs).

Mitigation efforts can be considered as having redundant sensors, modes of transportation, backup systems, and countermeasures within on-board diagnostics that differ in purpose from the diagnostics described earlier. When CAVs become dysfunctional, users need alternatives and seek out other modes of travel. Detection methods will be one type of mitigation measure that will employ selected vulnerabilities in the form of attack trees or graphs [1] for attacker moves and associated selected algorithms [2,3,4].

A second strategy is multi-modal connectivity that makes such choices for other modes of travel possible and will be a focus of the mitigation analysis. They can involve CAV and non-CAV modes. This portion of the research is an analysis of how multi-modal planning can increase resiliency when there is a cyberattack by increasing the types of travel, routing and scheduling available to users. Some of the kinds of multiple modes that will be investigated for selection as the basis for analyzing mitigation measures are multi-purpose, special purpose, substitutable, and those that have high connectivity with other modes.

Knowledge of individual mobility is an important prerequisite to developing strategies for the introduction of autonomous vehicles into transportation systems, since the use of CAVs may depend upon the suitability for the type of trip and competition from other travel modes already in place. Trustworthy information systems also play a role in the acceptability of CAVs. The purpose of this research component is to begin to build a knowledge base for the purpose of providing inputs for the development of algorithms for the introduction of CAVs in the form of multi-modal alternatives for mitigation measures and the incorporation of individual acceptance of and behavior with respect to CAVs. Specifically, the purpose is to provide mitigation measures so that CAV users can better trust the information they are obtaining.

Individual mobility is based on types of trips undertaken, the modes of travel for each, and where possible how the trips and modes are interconnected. Attributes of human behavior often shape these patterns and are also introduced. Obstacles to such travel patterns in terms of cyber-attacks are also briefly addressed. Definitions of multi-modal transportation have been summarized by [5] and multi-modality refers to connections between one or more forms of transportation that allow travelers to freely move among those different transportation forms. Research on modal connectivity by Zimmerman and colleagues has been presented under normal conditions [5] and in emergencies (e.g., [6,7] Zimmerman and Sherman [7] for example found that on September 11, 2001, travel mode changed over time and with the availability of transportation modes. At the outset they found that 93% used only a single mode. Later, to the ultimate destination, 55.3% used one mode, 26.7% used 2 modes, 13.1% used 3 modes, 4.2% used 4 modes and the rest used 5 or 6 modes. With respect to type of mode, at the outset, 92% walked but people gradually began to use other modes as they became available. Multi-modal connectivity in emergencies has been explored by other researchers [8, 9, 10] for evacuations and severe weather events some of which are analogous to security, as well as under normal conditions [11] that provide important background for and inputs to the research.

Subsection 4.2 Multi-modal Connections for Mitigation

The multi-modal portion of the research consists of several steps. The first is identifying generic structures of multi-modal connections organized by trip purpose and then how these would hypothetically change with the introduction of CAVs based on literature and cases about how such changes might occur. The second is public acceptance of CAVs in light of characteristics of CAV introduction.

4.2.1. Identification and Mapping of Multi-Modal Connections without and with CAVs by Trip

The first step in developing a framework for the hypothetical frequency of use of modal connections by trip purpose and readiness for CAV deployment or introduction is to understand travel behavior by trip purpose. Information about trip purpose is based on the U.S. Department of Transportation Federal

Highway Administration National Household Travel Survey (NHTS) [12]. The 2018 released NHTS survey provides such information for the U.S. and is shown in Table 1 for 2009 and 2017. Trip purpose components include (but are not limited to) direct trips for work, school, shopping (non-food and food), social purposes and recreation, and emergencies (not shown). Indirect trips are related to supplies to wholesale and retail establishments that ultimately provide products and services to consumers. The U.S. DOT FHWA [12: p. 18] data suggests that social and recreational trips dominate in 2017 accounting for over a quarter of household trips, however, in 2009, the U.S. DOT, FHWA and the U.S. DOT, BTS [13: p. 12] indicate the dominance of trips for family and personal errands.

Trip Purpose	2017	2009
Total All Purposes (number)	3,140	3,466
To/from Work (Percentages)	17.4	15.6
Work Related Business	1.6	3.1
Shopping	18.5	20.9
Other Family/Personal Errands	20.0	21.6
School/Church	10.9	9.6
School and Recreational	27.6	27.5
Other	4.1	1/8

Table 1. Distribution of Trip Purpose Type Over Time, 2009 and 2017

Source: [12] U.S. DOT, FHWA July 2018 Summary of Travel Trends: 2017 National Household Travel Survey, p. 18. https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf

The second step is to link modes of travel to trip purpose. The modes include (but are not limited to) Air, Water, Train (long distance to local), Truck, Vehicle (various occupancy options, buses, bikes (share, personal), Micro-mobility, and walking, recognizing that each of these categories have multiple subcategories based on size and technology. Data was drawn for example from the U.S. DOT National

Household Travel Survey that uses mode categories of private transport (e.g., private autos), public transit, walking and other. Another source of data is the Intermodal Passenger Connectivity Database (IPCD) is a national transportation facility database maintained by the U.S. DOT Research and Innovative Technology Administration (RITA), however, this does not generally include biking and walking. Table 2 below is an illustration of U.S. multimodal connectivity based on trip purpose and mode of travel from the U.S. 2018 NHTS databases:

	Mode(Percentages)			
Trip Purposes, 2017	Private	Public Transit	Walk	Other
All purposes	82.6	2.5	10.5	4.4
To/from Work	88.2	5.5	3.9	2.4
Work Related Business	80.1	3.4	8.4	8.0
Shopping; Errands	88.5	1.8	8.1	1.7
School/Church	70.5	2.5	10.3	16.7
Social and Recreational	77.1	1.6	18.1	3.3
Other	72.6	3.2	11.8	12.4

Table 2. Existing distribution of transportation mode by trip purpose, U.S.

Source: [12] U.S. DOT, FHWA July 2018 Summary of Travel Trends: 2017 National Household Travel Survey, p. 30-31. https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf.

Note: *The percent indicates the "percent of person trips by mode of transportation and trip purpose (millions). Constructed from https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf

The U.S. DOT FHWA [12: p. 23] data suggests that automobile travel dominates other modes, more in rural areas than urban areas given greater trip distances in rural areas; of the non-auto travel, walking dominates but micro-transportation modes are overtaking that. The dominance of auto travel has occurred historically for example from 1960-2010 [14: p. 29]. Combining trip purpose and mode, the

U.S. DOT FHWA [12: p. 30-31] data suggests that private vehicles are more frequently used for work trips (almost 90%) and relatively less frequently used (70-80%) for personal trips potentially reflecting trip length and variability.

The third step is to identify the linkages and where CAVs can supplement or substitute existing modes of travel by trip purpose. Criteria for such choices based on the perception literature include cost, performance, safety, trust, availability, etc. Table 2 above was reconstructed to reflect CAV substitution as an input to the models and will continue to be modified as modeling proceeds. This is illustrated in Table 3 below. Table 3 is just illustrative of a single scenario we use for the introduction of CAVs into existing modes of travel. It represents an increase in the percentage only for shopping and social/recreational trips, based on literature and case information. The purpose of the scenarios is to provide inputs into algorithms.

	Mode(Percentages)			
Trip Purposes, 2017	Private	Public Transit	Walk	Other
All purposes	82.6	2.5	10.5	4.4
To/from Work	88.2	5.5	3.9	2.4
Work Related Business	80.1	3.4	8.4	8.0
Shopping; Errands	95	2.5	2.5	0.0
School/Church	70.5	2.5	10.3	16.7
Social and Recreational	90.0	2.5	5.0	2.5
Other	72.6	3.2	11.8	12.4

Table 3. Hypothetical distribution of transportation mode by trip purpose assuming CAV substitution, U.S.

This distribution is generally consistent with an AAA [15] survey that found that almost three-quarters of drivers in the U.S. feared driving fully self-driving cars, however they would be relatively more accepting of the use of such vehicles if they were for short distances, for delivery services, and in relatively more confined spaces such as people mover systems.

Figure 8 portrays the hypothetical or potential introduction of CAVs into generic multi-modal trip patterns by trip purpose.

The integration of CAVs into existing multi-modal networks is a function of CAV system characteristics and requirements and user needs and the factors that contribute to it. CAV system needs include ([16]: the Existence of support facilities for CAVs, e.g., parking and refueling, Availability/Proximity of CAVs to other modes, Technology for and flexibility to connect with non-CAV modes, such as ridesharing and for-hire services, Communication capability to identify connections, Competition vs. compatibility with other travel modes, including congestion, travel time, cost, and complexity in assigning riders to vehicles [17: p. 5]. Overall supply logistics and strategy especially for emergencies [8, 9] and the attractiveness to passengers to use CAVs. The attractiveness dimension is addressed in the next section on user behavior.

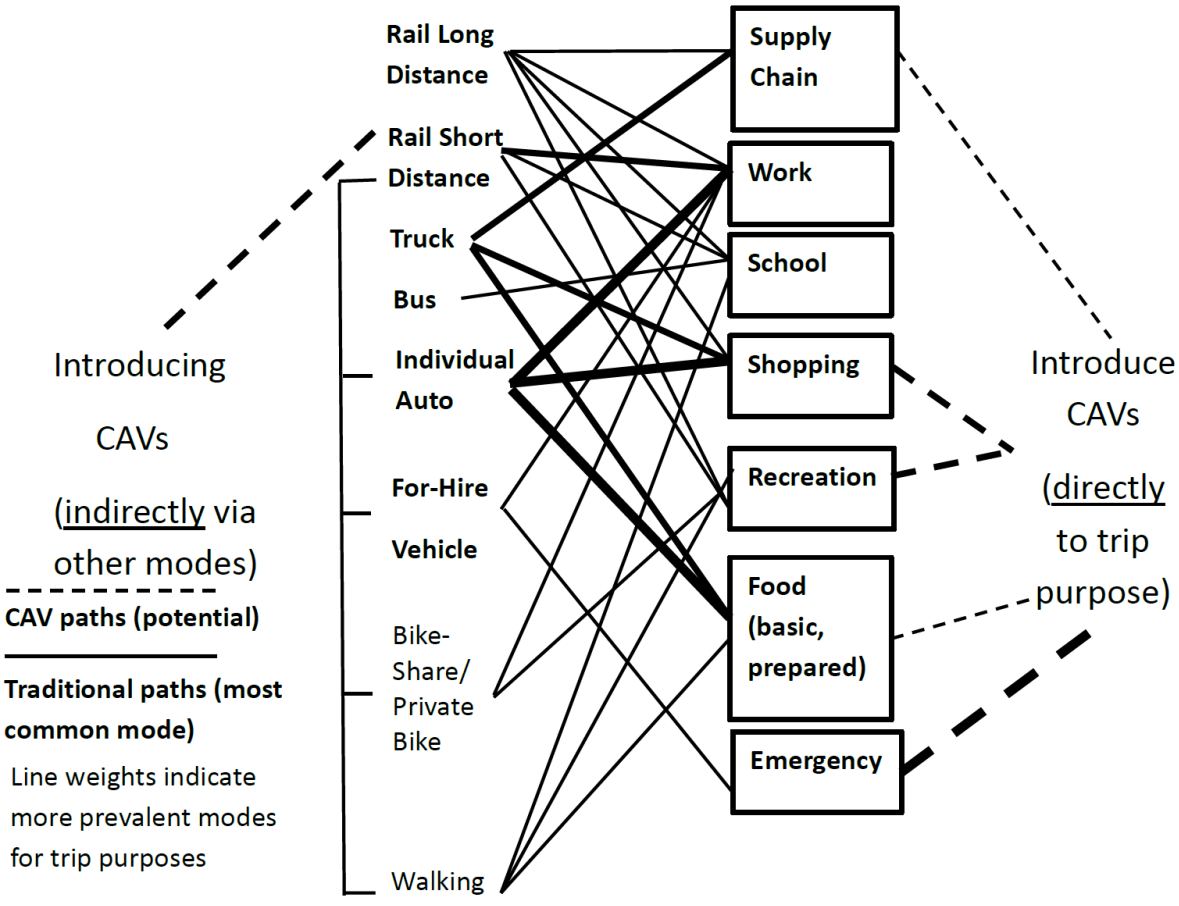


Figure 8. Relationships among Existing Multi-Modal Arrangements and Potential Arrangements with CAVs by Trip Purpose

4.2.2. Role of user behavior in CAV adoption

Insights about user attitudes and behavior toward CAVs were obtained through a literature review from a diverse set of journals and reports that specifically targets user behavior toward CAVs and the relationship to multi-modal connectivity. Much of the literature deals with cases or surveys, and the findings have been combined.

The attractiveness of CAVs to users in terms of their willingness to use them and pay for them is conceptually related to cost and willingness/ability of passengers to use CAVs, which in turn can be thought about in terms of safety/perceived fear [18], availability, convenience, and accessibility [19, 20]. Other factors contributing to perceived advantages pertain to travel time saving often associated with increased commuting time occurring for example between 2013 and 2017 [21]. Another consideration has been the provision of environmental benefits, such as the carbon footprint reductions associated with traditional automobile travel if number of trips are reduced [22,23].

Factors that adversely affect attitudes, perceptions, behavior and acceptance of CAVs tend to center around safety fears. Cyber security has been a major concern even in connection with traditional automobiles where multiple points of entry have been identified as potential cyber attack targets [24, 25, 26]. There are opportunities for protection, also involving synergies between physical security and cyber security for transportation systems in general [27].

These factors both positively and negatively are in turn based upon a long tradition in social psychology that are linked to individual attitudes toward control, certainty, the ability to assess probabilities and understand numerical information [28, 29, 30, summarized in 31].

Subsection 4.3 Summary

Trip purpose and mode usage by trip are important inputs into the evaluation of where CAVs could fit into multi-modal transportation. Trips that are flexible for example with respect to time, such as social/recreational trips, potentially benefit from CAVs. Very short trips regardless of trip purpose that are dominated by non-motorized transit or EV transportation are less likely to be substituted by CAVs given the ease of use of these other transit options relative to the greater complexity of CAV use. A fundamental aspect of the acceptability of CAV use is public knowledge, perception and behavior.

Sections 4.2 and 4.3 were prepared by Rae Zimmerman

References

[1] Dewri, R., Ray, I., Poolsappasit, N., and Whitley, D. "Optimal security hardening on attack tree models of networks: a cost-benefit analysis", International Journal of Information Security, 11: 3 (2012), pp.167-188.

[2] Zhu, Q. and Başar, T. "Game-theoretic approach to feedback-driven multi-stage moving target defense", In International Conference on Decision and Game Theory for Security, (2013), pp. 246-263, Springer.

[3] Miao, F., Zhu,Q., Pajic, M. and Pappas, G. J. "A hybrid stochastic game for secure control of cyber-physical systems", Automatica, 93 (2018), pp. 55-63.

[4] Mansheai, H., Zhu, Q., Alpcan, T., Basar, T. and Hubaux, J-P. "Game theory meets network security and privacy", ACM Computing surveys (CSUR), 45: 3 (2013), p. 24.

[5] Zimmerman, R., Restrepo, C.E., Sellers, J., Amirapu, A. and Pearson, T. R. Promoting Transportation Flexibility in Extreme Events through Multi-Modal Connectivity, U.S. Department of Transportation Region II Urban Transportation Research Center, New York, NY: NYU-Wagner, 2014. Final report available at: http://www.utrc2.org/sites/default/files/pubs/Final-NYU-Extreme-Events-Research-Report.pdf

[6] Restrepo, C.E., Zimmerman, R. and Joseph, R.A. with the assistance of Llopis Abella, J. 2017. "Public transit and mandatory evacuations prior to extreme weather events in New York City", U.S. Department of Transportation Region II Urban Transportation Research Center, New York, NY: NYU-Wagner, 2017. Final report. http://www.utrc2.org/sites/default/files/Final-Report-Public-Transit-Mandatory-Evacuations.pdf. For a brief summary see: https://www.transportation.gov/sites/dot.gov/files/docs/utc/April_2017_UTC_Spotlight.pdf

[7] Zimmerman, R. and Sherman, M. "To Leave An Area After Disaster: How Evacuees from the WTC Buildings Left the WTC Area Following the Attacks", Risk Analysis, 31: 5 (2011), pp. 787-804. First published online December 8, 2010. DOI:10.1111/j.1539-6924.2010.01537

[8] Ozguven, E.E., Ozbay, K., Iyer, S., Ryan A. G., and Whytlaw, J.A., Carnegie, "Multimodal logical architecture for emergency transportation toward better decision making in humanitarian logistics", 94th TRB Annual Conference (CD-ROM), Washington, D.C., Journal of Transportation Research Record (2015).

[9] Holguin-Veras, J. et al. "Emergency logistics issues affecting the response to Katrina: A synthesis and preliminary suggestions for improvement", Transportation Research Record, 2022 (2007), pp. 76-82

[10] Yang, X., Ban, X., and Mitchell, J. "Modeling multimodal transportation network emergency evacuation considering evacuees' cooperative behavior", Transportation Research, Part A (2017).

[11] Kaufman, S.M. and O'Connell, J. "Citi Bike: What current use and activity suggests for the future of the program". New York: New York University Rudin Center for Transportation Policy and Management, 2017.

[12] U.S. Department of Transportation, FHWA. "Summary of Travel Trends: 2017 National Household Travel Survey", 2018. https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf

[13] U.S. Department of Transportation, Bureau of Transportation Statistics "Passenger Travel Facts and Figures 2015", 2015. http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/PTFF_Complete.pdf

[14] U.S. Environmental Protection Agency "Our Built and Natural Environments: A Technical Review of the Interactions among Land Use, Transportation, and Environmental Quality", Second edition. Washington, D.C.: EPA, 2013. http://www.epa.gov/dced/pdf/b-and-n/b-and-n-EPA-231K13001.pdf.

[15] American Automobile Association "2019 Fact Sheet Automated Vehicle Survey —Phase IV" 2019. https://newsroom.aaa.com/2019/03/americans-fear-self-driving-cars-survey/

[16] Zimmerman, R. "Challenges of multimodal networks interconnecting transportation technologies: CAV and other mode innovations", 2019 Transport-Tech Summit, New York, NY, 2019.

[17] National Academies of Sciences, Engineering, and Medicine (NASEM) "Shared Mobility and the Transformation of Public Transit", Washington, DC: The National Academies Press, 2016. https://doi.org/10.17226/23578

[18] Litman, T. Autonomous vehicle implementation predictions implications for transport planning. Victoria, British Columbia: Victoria Transport Policy Institute, 2019.

[19] Nordhoff, S., van Arem, B. and Happee, R. "Conceptual model to explain, predict, and improve user acceptance of driverless podlike vehicles, Transportation Research Record: Journal of the Transportation Research Board, 2602, 2016. Transportation Research Board, Washington, D.C., pp. 60–67. DOI: 10.3141/2602-08

[20] Fraedrich, E. and Lenz, B. Societal and individual acceptance of autonomous driving. In: Maurer, M., Gerdes, J.C., Lenz, B., Winner, H. (Eds.) Autonomous Driving. Technical, Legal and Social Aspects, Springer, 2016, pp. 621-640. DOI 10.1007/978-3-662-48847-8_29

[21] U.S. Census Bureau, American Community Survey 2010.

[22] Brown A., Gonder J., and Repac, B. 2018. An analysis of possible energy impacts of automated vehicles in: G. Meyer and S. Beiker (eds.), Road Vehicle Automation, Lecture Notes in Mobility, Lecture Notes in Mobility, 2018. Springer, Cham. https://doi.org/10.1007/978-3-319-05990-7_13

[23] Zimmerman, R. and Zhu, Q. 2020. Abstract: "Attitudes and perceptions potentially associated with behavior toward autonomous vehicles," International Symposium for Sustainable Systems and Technology (ISSST) 2020, June 8-12, 2020, Pittsburgh, PA (online)

[24] McAfee "Caution: Malware ahead:, 2011. http://www.mcafee.com/us/resources/reports/rp-caution-malware-ahead.pdf.

[25] Zimmerman, R. Transport, the Environment and Security. London, UK: Routledge, 2012.

[26] Zimmerman, R. "The cyber and critical infrastructures nexus: interdependencies, dependencies and their impacts on public services" White Paper, NYU Center for Cybersecurity, 2017. http://cyber.nyu.edu/wp-content/uploads/2017/12/Zimmerman-122717updateEdited-White-Paper-Cyber-and-Critical-Infrastructures.pdf

[27] Zimmerman, R. and Dinning, M. G. "Benefits and Needs for an Integrated Approach to Cyber-Physical Security for Transportation", In Transportation Systems Resilience: Preparation, Recovery, and Adaptation, Transportation Research Circular E-C226, Transportation Systems Resilience Section, Standing Committee on the Logistics of Disaster Response and Business Continuity, Standing Committee on Emergency Evacuations, Standing Committee on Critical Transportation Infrastructure Protection, Transportation Research Board, Washington, DC: National Academies Transportation Research Board, 2017, pp. 15-21. http://onlinepubs.trb.org/onlinepubs/circulars/ec226.pdf

[28] Slovic, P., Fischhoff, B. and Lichtenstein, S. "Facts vs. fears: Understanding perceived risk", In: Judgment under uncertainty: Heuristics and biases. Edited by D. Kahneman, P. Slovic, and A. Tversky, Cambridge, Eng.: Cambridge U. Press, 1982, 481-482.

[29] Slovic, P., Fischhoff, B. and Lichtenstein, S. "Rating the risks", in The Perception of Risk, edited by P. Slovic. Sterling, VA and London, UK: Earthscan, 2000, pp. 104-120.

[30] Peters, E., Hibbard, J., Slovic, P. and Dieckmann, N. "Chapter 20, Numeracy skill and the communication, comprehension and use of risk-benefit information", in P. Slovic, The Feeling of Risk, London and Washington, DC: Earthscan, 2010, pp. 345-352.

[31] Zimmerman, R. "Human behavioral factors that shape urban physical infrastructure services, in Proceedings from EDRA 50: Sustainable urban environments, edited by A. Beth, R. Wener, B. Yoon, R. A. Rae, and J. Morris. Brooklyn, NY: Environmental Design Research Association, 2019.

https://cuny.manifoldapp.org/read/human-behavioral-factors/
https://cuny.manifoldapp.org/read/human-behavioral-factors/section/c3a3b0d7-a9e2-4686-af8c-6e41b5dcb52f.

Section 5: Final Report and Publications

C2Smart Year 1 Publication and Presentation Listing

Documents are listed approximately in inverse chronological order

Publications

T. Kieras, M. J. Farooq, Q. Zhu (2020) Holistic Risk Analysis of Hierarchical Supply Chain Threats in Connected and Autonomous Vehicles, 2020 IEEE World Forum on Internet of Things, New Orleans, USA, under review.

S. Zhou, J. Pang, and Q. Zhu (2020) MUSIC: A High-Capacity Low-Waiting Time Intersection Management System for Connected Autonomous Vehicles, under review.

T. Zhang and Q. Zhu (2020) Differentially Private Collaborative Intrusion Detection Systems for VANETs, arXiv:2005.00703

R. Zimmerman (2020) Human-Made Disasters: Electric Power and Transit Linked Outages. In: Shapiro L., Maras MH. (eds) Encyclopedia of Security and Emergency Management. Springer, Cham. https://doi.org/10.1007/978-3-319-69891-5; https://link.springer.com/referenceworkentry/10.1007/978-3-319-69891-5_291-1 Copyright © 2020, Springer Nature Switzerland AG.

R. Zimmerman and C.E. Restrepo (2019) Physical Security: Exterior Application. In: Shapiro L., Maras MH. (eds) Encyclopedia of Security and Emergency Management. Springer, Cham. https://doi.org/10.1007/978-3-319-69891-5; https://link.springer.com/referenceworkentry/10.1007/978-3-319-69891-5_221-1 Copyright © 2020, Springer Nature Switzerland AG.

R. Zimmerman (2020 forthcoming) "The Interdependencies of Infrastructure" [revised title] in Urban Infrastructure: Interdisciplinary Perspectives from History and the Social Sciences, edited by J. Soffer, J. Heathcott, and R. Zimmerman. Pittsburgh, PA: University of Pittsburgh Press (in contract), in preparation, forthcoming. Paper Outline

Presentations

May 15, 2020. R. Zimmerman, The integration of big data for impacts of COVID19 decisions points on public service resilience, in session titled, Data Driven Disaster Planning, INFORMS 2020, invited presentation.

April 24, 2020. R. Zimmerman, Abstract Submission for Conference Presentation, Association for Public Policy Analysis and Management (APPAM) Annual Meeting November 12-14, 2020, Washington, DC, Washington, D.C. "Human behavior and the transformation of public infrastructure services under extreme conditions." Topic Area 2. Innovations in Science and Technology

March 10, 2020. R. Zimmerman, Abstract Submission for Conference Presentation, Association of Collegiate Schools of Planning (ACSP) 60th Annual Conference, November 4-8, 2020, Toronto, Ontario Canada (online likely), "Incorporating Public Acceptance into Planning Decisions for Autonomous Vehicles," Transportation group.

March 6, 2020. Rae Zimmerman and Quanyan Zhu, "Integrating New Transportation Technologies into Existing Travel Modes: The Case of Autonomous Vehicles," Transportation session. NYU Faculty Urban Research Day. New York: NYU. Abstract.

March 6, 2020. Quanyan Zhu and Rae Zimmerman, "The Challenge of Achieving Urban Resilience with Interdependent Infrastructures and Extreme Events", Safety and Resilience session. NYU Faculty Urban Research Day. New York: NYU

March 5, 2020 acceptance, January 25, 2020 submission R. Zimmerman and Quanyan Zhu, Abstract Submission and Acceptance for Conference Presentation, International Symposium for Sustainable Systems and Technology (ISSST) 2020, June 8-12, 2020, Pittsburgh, PA (online likely), "Attitudes and Perceptions Potentially Associated with Behavior Toward Autonomous Vehicles," Oral presentation – webinar (online), June 23, 2020. Abstract.

January 15, 2020. Transportation Research Board, Zimmerman presentation titled "The Challenge of Physical Security Adaptation for Multi-Hazards for Resilient Transportation" (P20-20144). Presentation Slides.

January 13, 2020: R. Zimmerman, co-chair of Physical Security subcommittee (ABR10(3)) January 13, 2019 Monday, Marriott Marquis hotel, 6:00 PM-7:30 PM Liberty Salon P (Level 4 (M4)) R. Zimmerman, "Managing Adverse Human Behavior with Respect to Transportation Infrastructure Security," Lightning talk.

December 10, 2019. R. Zimmerman, "A data-driven framework for user, provider, and community behavior toward infrastructure services risks," Society for Risk Analysis 2019 annual meeting, Arlington, VA. Abstract.

November 1, 2019. R. Zimmerman, "Challenges of Multimodal Networks Interconnecting Transportation Technologies: CAV and other Mode Innovations" 2019 Transport-Tech Summit, New York, NY.

June 24, 2019. R. Zimmerman, closing remarks for Sustainable Urban Subsurface Systems Workshop, NSF funded workshop, "DIG Around: Documenting Infrastructure and the Ground Conditions Around It" Brooklyn, NY; June 24-25, 2019, led by the NYU Center for Urban Science and Progress (CUSP). NSF #1929923. Invited Organizing Committee Member and commentary.

May 23, 2019. R. Zimmerman, "Human Behavioral Factors that Shape Urban Physical Infrastructure Services." Paper presentation at the Environmental Design Research Association (EDRA): EDRA50 Sustainable Urban Environments: Research, Design and Planning for the Next 50 Years. Abstract link: https://edra.confex.com/edra/EDRA50/meetingapp.cgi/Paper/6734

Section 6: Development of Outreach Plan

Subsection 6.1 Outreach objectives and approach

Outreach is an important part of engaging potential users and other interested parties in the project and its outcomes. The outreach plan consists of identifying members of the industry, academic, professional, and educational communities who can potentially benefit from the research, describing how they were identified and engaged over the course of the project, and developing a process to provide research outputs for such communities in various ways. Outreach efforts are expected to be ongoing over the course of the research.

Subsection 6.2 Educational Outreach

Outreach to educational communities consists of involving students in the research in a number of ways. Students are considered an important part of the educational outreach for the project. First, a few students are hired to work directly on the project. Two graduate students are hired to work on the project and a number of undergraduates are involved as well though VIP projects. Second, outreach to a broader set of students is provided through academic courses and other training venues such as summer undergraduate research internships, and projects within courses developed by Professor Zhu.

For student and workforce development, Zhu has developed new courses, tutorials and workshops that aim to train the next generation in the multidisciplinary technology frontier that intersects cybersecurity, critical infrastructure, and social behaviors.

Subsection 6.2 Industry Academic and Professional Community Outreach

The plan for outreach to industry, academic and professional communities consists of (1) identification of selected members and the method of identification and (2) ways in which these entities are planned to be used and their potential engagement.

6.3.1 Identification and method of identification

Organizations or individuals who potentially benefit the most from the research are first identified according to the researchers' initial judgement. The process consists of contacts in the investigators' immediate network and those that have come to light in literature and cases. These will be organized in terms of research components, e.g., algorithms analysis and mitigation plans. When an initial set of contacts is identified, these contacts are requested to suggest others in their network.

6.3.2 Ways in which organizations and individuals are planned to be used and potential engagement

Organizations and individuals in industry, academia, and various professions will be used to provide inputs into the findings of the project as they are developed and to provide contact points for the dissemination of the research. Engagement is planned through a variety of venues that will focus on one on one contacts rather than group meetings organized by the project. The researchers plan to take advantage of workshops and other events organized by other research groups and presentations at relevant conferences attended by both academicians and professionals. This approach of using existing platforms was considered effective to maximize the research exposure to these groups rather than holding separate events. While these activities consist of a diverse set of organizations, many provide rich contacts to the autonomous vehicle community though they may not directly be involved in the field.

One set of events includes a workshop held in connection with colleagues Another was a workshop held in October 2018 by C2Smart on autonomous vehicles. A third workshop was convened on June 20, 2019 by various universities under the auspices of CCNY, NYU, and other institutions in connection with a partnership for a potential Engineering Research Center. A fourth and major ongoing venue is major national conferences and meetings as well as local events where the PI and co-PI routinely present their work.

The Transportation Research Board annual conferences and committee and subcommittee activities are an important venue for the research. Zimmerman has served a nine-year maximum term as an invited appointed member of TRB's ABR10 (Critical Transportation Infrastructure Protection Committee) and was appointed co-chair of the committee's Physical Security subcommittee. During that time, she

organized conference sessions, presented formal papers, and lightning talks. She also attended meetings and conference calls of the cyber security subcommittee that provide important contacts and publishing venues. The TRB connection has provided a number of contacts for the research including the U.S. DOT Volpe Center, the FHWA, and NYC-based transportation agencies such as the MTA, the Port Authority of NY and NJ, the NYC DOT, and NYMTC.

In addition, other national and international society venues that the researchers are routinely in contact with include the Society for Risk Analysis, Resilience Week, IEEE, INFORMS, and the Association of Collegiate Schools of Planning. Various professionals listed under the partner section, including the new U.S. DHS CISA, are anticipated to be integrated into outreach efforts. Professor Zhu has developed contacts through a number of events he has organized after this research commenced such as the 9th International Conference on NETwork Games, COntrol and OPtimisation (NETGCOOP) in 2018, the 5th International Conference on Artificial Intelligence and Security (ICAIS 2019) in 2019, and 2020 IEEE Workshop on Information Forensics and Security (WIFS).

Included as part of the plan is outreach to formal professional associations engaged in security-related transportation research such as ISACs and NIST.

Zhu and Zimmerman plan outreach to the research communities by giving talks or attending events at Historically Black Colleges and Universities (HBCU) schools, such as John Jay College, student-organized events, and giving tutorials and online courses to students and junior researchers.

Outreach includes drawing upon a number of contacts in the areas covered in the proposal, including some of the professionals in those areas that C2SMART tapped in their recent symposium (October 2018). In line with the mitigation measures, partnerships are sought for example with the MTA (sustainability and resilience area) in addition to others mentioned above.

The outreach plan includes researchers planning to participate in the workshops at ACM CCS workshop on Cyber-Physical Systems Security & Privacy (CPS-SPC), CPSWeek Workshop SCAV: Safe Control of Autonomous Vehicles and the local NYU CSAW events every year.

Section 7: Technology Transfer, Data Repository, Guide and Demonstration Online

The research will generate selected research products prepared for readiness for transfer to potential users in the form of data and strategies for CAV characterization and mitigation of adverse impacts. One set includes algorithms and codes and analytical methods that can be broadly applied to identify V2V V2I and vehicular security issues and mitigate them by reducing the impacts. A second set is generated from the mitigation plan including technological solutions as well as planning solutions. These not only benefit the users but also the infrastructure planners and technology developers. Potential users are those identified in the outreach plan.