Driver Identification and Authentication with Active Behavior Modeling

Angela Burton *, Tapan Parikh†, Shannon Mascarenhas†, Jue Zhang†, Jonathan Voris†, N. Sertac Artan ‡ and Wenjia Li†
*Department of Electrical Engineering and Computer Science, Vanderbilt University
Email: angela.j.burton@vanderbilt.edu
†Department of Computer Science, New York Institute of Technology
Email: tparikh@nyit.edu, smascare@nyit.edu, jzhang47@nyit.edu, jvoris@nyit.edu, wli20@nyit.edu
‡Department of Electrical and Computer Engineering, New York Institute of Technology
Email: nartan@nyit.edu

***Abstract*—The legitimate driver of a vehicle traditionally gains authorization to access their vehicle via tokens such as ignition keys, some modern versions of which feature RFID tags. However, this token-based approach is not capable of detecting all instances of vehicle misuse. Technology trends have allowed for affordable and efficient collection of various sensor data in real time from the vehicle, its surroundings, and devices carried by the driver, such as smartphones. In this paper, we propose to use this sensory data to actively identify and authenticate the driver of a vehicle by determining characteristics which uniquely categorize individuals' driving behavior. Our approach is capable of continuously authenticating a driver throughout a driving session, as opposed to alternative approaches which are either performed offline or as a session starts. This means our modeling approach can be used to detect mid-session driving attacks, such as carjacking, which are beyond the scope of alternative driver authentication solutions. A simulated driving environment was used to collect sensory data of driver habits including steering wheel position and pedal pressure. These features are classified using a Support Vector Machine (SVM) learning algorithm. Our pilot study with 10 human subjects shows that we can use various aspects of how a vehicle is operated to successfully identify a driver under 2.5 minutes with a 95% confidence interval and with at most one false positive per driving day.**

I. Introduction

Ignition keys have served as authentication tokens for vehicle drivers for decades. More recently, traditional keys based on physical shape have been augmented with embedded Radio Frequency Identification (RFID) tokens to provide an additional layer of protection against theft. Unfortunately, such keys are susceptible to theft, cloning, forgery, and relay attacks. However, RFID enabled steering columns represent only a small portion of the sensing hardware available on modern vehicles.

Traditionally, measuring driver habits involved conducting costly traffic surveys which take a large amount of time and human effort yet yielded results with limited accuracy [1]–[3]. Recently, however, several technological trends have converged to allow affordable and efficient collection of driver data [4], [5]. The cost and availability of wireless communication and sensing hardware has allowed for easy collection of data, often in real time via ubiquitous devices installed in vehicles or worn by drivers. This has enabled a variety of potential applications, including more accurate pricing determinations for insurance (Pay as you drive (PAYD) or Usage-Based Insurance (UBI)) [6], [7] and finer grained traffic planning for improved public safety [8].

This research seeks to solve the problem of unsafe and untrustworthy transportation systems caused by vehicle misuse by authenticating drivers according to the manner in which a vehicle is operated. To this end, we conducted a study with 10 human subjects to assess the efficacy of using the data collected from the vehicle sensors on identifying the driver.

In this paper we propose to authenticate drivers based on a variety of data that is available via common onboard vehicular sensors and systems. There are a variety of stakeholders involved in the operation of transportation systems for which a more thorough guarantee of a driver's identity would be of interest. For example, municipal governments may wish to ensure that buses are being operated by a predetermined employee. Similarly, car sharing service providers may want to confirm that a member has picked up the correct vehicle, and owners of taxi fleets may wish to ensure their vehicles have not been operated without permission. Insurance providers may wish to verify that only drivers listed on a particular policy are allowed access to a covered car. Finally, recognition that a vehicle is being operated by someone other than the vehicle's typical owner may allow for advanced notice in the event of vehicle theft.

The remainder of this paper is organized as follows. Section II summarizes pertinent related work. Section III introduces our threat model, discusses the design and experimental setup of our human subject study, and discusses our approach to the problem of driver behavior modeling. Section IV presents the outcome of our study. Section V concludes the paper.

II. Related Work

Vehicle-based performance technologies infer driver behavior by monitoring car systems such as lane deviation, steering or speed variability [9]. Such systems are critical to identify and avoid driver sleepiness, which is associated with around 20% of serious car injuries [10]. Jensen *et al.* proposed mathematical strategies to classify drivers into various levels of aggressiveness using vehicle information such as vehicle

978-3-901882-85-2

speed, engine speed, mass air flow rate, coolant temperature, and throttle [11].

In recent work, such systems are suggested for driver classification. Salemi proposed [12] to use vehicle information such as the vehicle speed, engine speed, breaking speed, and gyro data to profile the driver behavior. They use this profile to authenticate a driver to avoid car theft. However, such systems require access to a vehicle's engine via a debug interface, known as On-Board Diagnostics (OBD-II), which may introduce a security risk by providing a vector for attack through influencing vehicle control variables such as steering, braking, and accelerating [13]. GPS tracking, which may violate a driver's privacy and lead to possible legal and social implications [14], has been criticized for issues such as insurance pricing manipulation. Researchers recently showed that such systems can pinpoint a driver's destination even without GPS [15].

There exist sensing technologies for directly monitoring driver fatigue using cameras placed on the dashboard [16] or via biological modalities such as electroencephalography (EEG) and electrooculography (EOG). These technologies are cumbersome and intrusive, though, thus limiting their widespread use. Following the advances in wearable technology, recent work started addressing some of the aforementioned challenges using wearable devices for driver identification. Karatas *et al.* proposed using wrist-worn wearables for tracking drivers and identifying drivers by profiling the drivers' steering behavior [17]. This and some other driver identification schemes [18], [19] have been presented in previous work, but the approach presented in this paper is the first to the authors' knowledge to be capable of detecting mid-driving session attacks such as carjacking.

Vehicular sensing applications which respect occupant privacy is an emergent line of work. For example, Troncoso *et al.* studied how PAYD insurance model can still be feasible by revealing only the account information required for billing purposes rather than all aspects of a driver's activity [20]. This is achieved by a privacy-preserving mechanism, which performs calculations locally, and sends only the aggregated data to the insurance company. Checkoway *et al.* carried out a formal analysis of the ways in which connecting devices to a vehicle can increase its attack surface [21]. Authenticating users prior to driving activity is another emerging research field. In this context, beyond using car keys or tokens for access control to the car, to allow the vehicle activation drivers are required to be authenticated by verifying their posture [22], [23] or fingerprints [24].

III. Method

A. Threat Model

There are clearly differences between authenticating a driver to a vehicle and user authentication in traditional contexts which impact adversarial capabilities. Whereas typical users are fairly stationary and may walk away from a session at any time, driving a vehicle is a high-speed operation which cannot be terminated abruptly. We thus assume that the same driver is in control of the vehicle throughout a session for practical considerations, although a strength of our model is its capability to detect changes in behavior even within a single driving session.

Instead, we concentrate on attack situations in which an unauthorized individual has somehow gained access to a vehicle by bypassing existing authentication mechanisms. We envision a variety of potential cases in which this could occur, involving different combinations of stakeholders:

- A single-owner vehicle is stolen and driven away. Perhaps the car is broken into by stealing a key, immobiliser, or other access token. Vehicles can also be "hot-wired" by circumventing the ignition interlock. Alternatively, a careless owner may have simply left his or her vehicle unattended and running; this parallels "lunchtime" attacks against traditional computer systems [25].
- In many jurisdictions only insured drivers are permitted to operate certain vehicles; operation of a vehicle by a party not authorized by the insurance agency constitutes a legal violation [26].
- Operators of transportation services, such as taxi and bus companies, may wish to verify that certain drivers are completing the correct routes for compensation and liability purposes.
- Car rental agencies, and other car sharing providers, frequently stipulate that only specific customers, clients, or members may operate a vehicle during a particular booking.

We assume an adversary with no special familiarity to the driving patterns of other users; the impact of driving knowledge on the ability of an attacker to emulate driving patterns will be pursued as future work.

B. Experimental Study Design

For our study, we used the OpenDS 3.5 driving simulator, which is based on the JMonkey Engine [27]. We used a Logitech G27 Steering Wheel in place of a traditional mouse and keyboard so that users could perform their driving task in a more realistic manner.

Our study participants were asked to complete a specific driving task using a scenario we designed using the OpenDS driving simulation software [27]. During our study, each of the 10 test subjects were asked to complete four 5 minute laps with the OpenDS simulation, and consequently we collected approximately 20 minutes of driving data per subject. The simulation is set to have light traffic, frequent traffic light changes, pedestrians, and road signs in order to simulate a realistic driving experience.

During each session with the simulator, five driving parameters are collected from OpenDS at approximately 40 millisecond intervals. These parameters are vehicle position (X, Y, and Z coordinates), speed (km/h), steering wheel position, gas pedal position, and brake pedal position.

C. Data Modeling and Analysis

We chose five features as potentially discriminative characteristics extracted from measurements to capture each subject's unique driving pattern: (1) Euclidean distance traveled, (2) average vehicle speed, (3) the standard deviation of the steering wheel position, (4) the average change of brake pedal position, and (5) the average change of gas pedal position. These features were selected for a combination of practical and theoretical considerations.

The OpenDS driving simulation software's logging functionality allowed easy access to low level driver tracking details from which each of these features could be derived. Additionally, we felt that these features would be good candidates for capturing driving activity because they covered a wide range of the various controls one must utilize in order to drive proficiently. Lastly, variations of these modalities had previously been considered for different but related modeling tasks [11], [12], [17]. We also included the vehicle's location coordinates and rotation in order to provide a basis for comparison with our derived features.

We applied a variety of different machine learning algorithms to our collected feature set in order to assess their ability to discern between individuals as they operated a vehicle. We implemented Matlab scripts to apply 3 different supervised learning algorithms to our data: Decision Trees, Support Vector Machine (SVM), and k-Nearest Neighbor (kNN). We also attempted to apply a boosting to increase our classification accuracy: instead of using all features for classification, various subsets of features are used and classification is determined by which grouping is indicated by a majority of the learners. In the case of k-Nearest Neighbor, this is referred to as the random subspace method, while for decision trees, this results in an approach known as Random Forests.

We plotted Receiver Operator Characteristic (ROC) and Detection Error Tradeoff (DET) curves to provide a fair comparison between these disparate classification techniques. An ROC curve is a plot of a classifier's true positive rate, or sensitivity, against it's false positive rate as the threshold for classification is altered. Values to the lower left of the ROC curve represent more *conservative* threshold values, with less false positives (*i.e.,* false alarms about an authentic driver's identity) but also less true positives (*i.e.,* a less successful unauthorized driver detection rate). The upper right of the ROC curve, on the other hand, shows less *conservative* thresholds where attacker detection is maximized at the cost of increased false positives. Because the goal of driver classification is to maximize the true positive rate of detection while minimizing the number of false alarms raised during regular driving activities, the goal is to maximize the area under the ROC curve (AUC).

DET curves are very similar to ROC curves in that both plot classifier performance as a function of threshold adjustment. A DET curve plots a classifier's true positive rate against its false positive rate, however, while a DET curve instead plots a classifier's false rejection rate against its false positive rate. DET curves are useful for visualizing the relationship between these error rates. The point at which both error rates equal each other is known as the Equal Error Rate (EER).

We conducted our SVM classification tests with the C-support vector classification (C-SVC) training algorithm. Using logistic models with a complexity parameter, which is used to control the number of support vectors used to form class boundaries, of ($C = 10.0$), resulted in the best achievable classification performance.

D. Feature Analysis

To measure how well these features capture patterns specific to a driver, we utilize Fisher's separation function [28]. This function, which is maximized to classify data in Linear discriminant analysis, is the ratio between matrices representing a feature's scatter within a class and across classes. Due to the proportionality between scatter and covariance, a Fisher score can be expressed as the ratio between the within-class variance and between-class variance of a feature [29]:

$$s = \frac{\sum_{i=1}^{c} n_i (u_i - u)^2}{\sum_{i=1}^{c} n_i \sigma_i} \tag{1}$$

where the numerator is the between-class variance for feature n and the denominator is the within-class variance for feature n.

E. One-Class Modeling

We explored the application of multiclass modeling processes to the task of driver identification in order to perform a comparison of alternative modeling techniques. In practice, however, a particular driver's vehicle would not have access to information regarding how other drivers operate their vehicles. Furthermore, even if this information was available, it would be very difficult to scale to all users in a busy driving area. For this reason, one class models, which require only positive samples of an authentic driver's behavior patterns, are much better suited to the context of driver authentication.

To see how a one-class model would perform with respect to our driving features, we applied a one-class Support Vector Machine (oc-SVM) to our data to create a separate model for each user. Each user's model was trained with 80% of their driving samples, while the remaining 20% of driving logs were reserved for testing. Each user's model was trained only using their driving data, but driving data from all ten subjects was used to test the classification accuracy.

F. Time To Detection

Though the previous section's analysis is useful for the sake of comparing our approach to different modeling techniques and behavioral modalities, it is also instructive to consider the operational considerations of our approach. Specifically, we would like to know how long it would take to identify someone who has attempted to illicitly operate a car. The time to detection (TTD) is a function of the true positive rate of our driver modeling system, which is in turn proportional to our

model's false positive rate. It is thus necessary to fix a false positive rate before determining our model's TTD.

Once a desired operational false positive rate is selected, it can be used to determine a target per sample rate of false positives. The resulting false positive rate of the classifier can then be used to determine the true positive rate of the classifier at that specific threshold value. Finally, the per sample true positive rate can be used to determine how many samples, and thus how much observation time, is required to achieve a desired detection confidence rate.

IV. Results

A. Model Comparison

We plotted the Receiver Operator Characteristic (ROC) curves to provide a fair comparison between these disparate classification techniques. Since ROC curves express binary classification information, this was performed on a user by user basis, *i.e.,* user 1 was classified against users 2 through 10, then user 2 was compared to user 1 and users 3 through 10, etc.

More specifically, we conducted our SVM classification tests with the C-support vector classification (C-SVC) training algorithm. Using logistic models with a complexity parameter, which is used to control the number of support vectors used to form class boundaries, of ($C = 10.0$), resulted in the best achievable classification performance.

Figure 1 shows ROC curves which result from multiclass SVM classification for all ten participants. In addition to the ROC curves, Figure 1 also provides AUC values for each study participant that resulted from training an SVM on our driving features with a Polynomial kernel and applying 10-fold cross validation; the average AUC across all users is 0.8138.

Figure 2 presents a DET curve for multiclass SVM classification averaged across all users. Our multiclass driver detection SVM was capable of authenticating drivers with an EER of 24.9%.

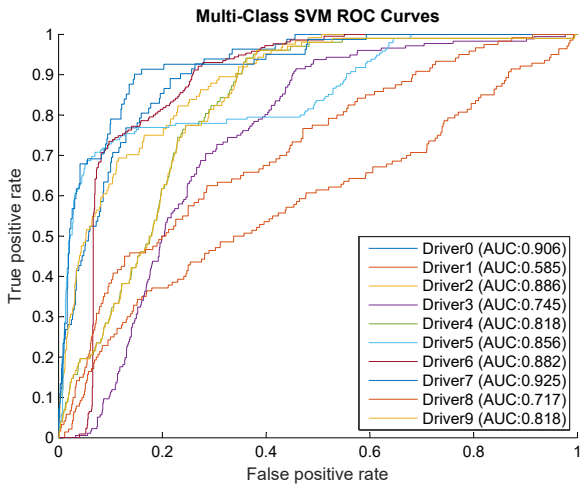

Fig. 1. ROC Curves for Multi-Class SVM Classification of All Study Participants.

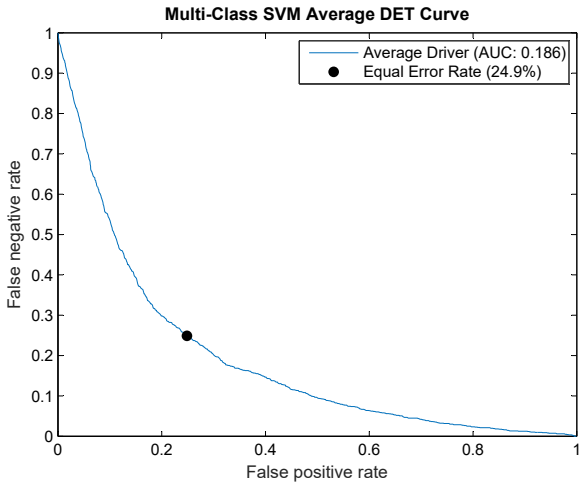

Fig. 2. Average DET Curve for Multi-Class SVM Classification.

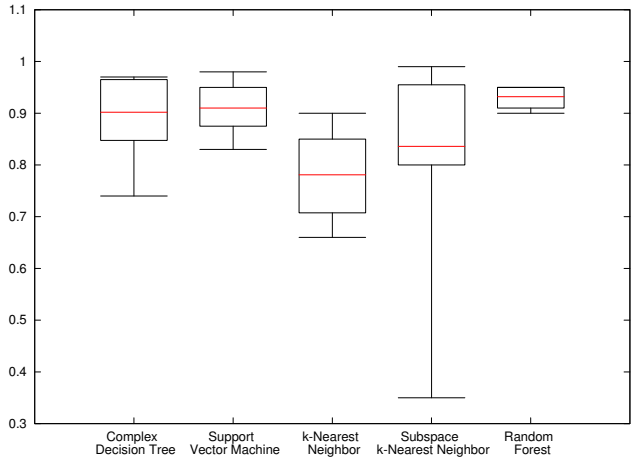

Fig. 3. Multiclass Model AUC Comparison.

These per-user AUC values were averaged together to produce an AUC for each classifier, which is displayed in Figure 3 as a box-and-whisker plot. Decision trees displayed a similar classification performance to SVMs on average, resulting in AUC values of 0.902 and 0.91 respectively. However, decision trees also fell into the lower quartile for a larger portion of users. The kNN approach displayed the worst overall performance, with an average AUC of 0.781. Though boosting did increase the average AUC to 0.8 for kNN, it also resulted in very poor performance for some users, with an AUC value as low as 0.35. Though boosting decision trees to produce a random forest ensemble learner increased the average AUC to 0.932, we believe this gain may be due in part to overfitting on our relatively small sample size.

B. Feature Analysis

The Fisher scores for both the average values of our raw driving data and our derived features are shown in Table I. The features are listed in ascending order by their Fisher

TABLE I
FISHER SCORES FOR DRIVING FEATURES.

Feature	Fisher Score	Classification Contribution
Average Change in Accelerator Pressure	0.122	3.84%
Distance Traveled	0.101	0.23%
Average Speed	0.082	0.26%
Average Change in Brake Pressure	0.052	1.76%
Standard Deviation of Steering Position	0.039	0.60%
Average X Axis Position	0.037	0.46%
Average Z Axis Position	0.022	1.32%
Average Y Axis Position	0.020	0.00%
Average Z Axis Rotation	0.019	0.00%
Average Y Axis Rotation	0.018	-0.46%
Average X Axis Rotation	0.017	-0.03%
Average W Axis Rotation	0.014	0.07%

score. These scores capture the ratio of between-class and within class variance, which essentially means that higher ranked features are more consistent for a particular driver over time, and more unique between different drivers. From Table I, it is easy to see that our derived values have more discriminative power than the "raw" rotational and coordinate values collected from the simulator; recall that the coordinates are roughly equivalent to geolocation information.

The third column of Table I, labeled "Classification Contribution," contains another measurement of the suitability of each feature to the task of driver modeling. This value is obtained by removing the feature from our SVM modeling process and observing the new true positive classification accuracy. The new TPR is subtracted from the original to obtain the classification contribution.

As shown in Table I, the classification contributions are correlated with Fisher values, with higher valued features having larger contributions to the overall modeling process. A notable exception are the Distance Traveled and Average Speed. The reason why SVM modeling retains its classification accuracy when either of these features is removed is due to the fact that they are highly correlated, thus removing one or the other only removes a small amount of information from our models due to the redundancy in these features. Some measurements, particularly the Y axis position, which represents elevation, are consistent across all users. The Fisher score and classification contributions confirm that these features are not discriminative. Including the Y and W rotational axes in our model even turned out to be detrimental to classification. We believe the reason for these features to have non-zero Fisher scores is due to noise in the underlying data introduced by very small variations in the data logged by the simulation.

C. One-Class Modeling

Figure 4 shows the ROC curves achieved for each driver using the per-user oc-SVM modeling process, while Figure 5 displays the average DET curve for all users.

The oc-SVM achieved an average AUC value of 0.9219 and an EER of 14.7%, which represents an improvement over multi-class modeling in terms of both metrics. This is due in part to the fact that the one-class modeling process is asking

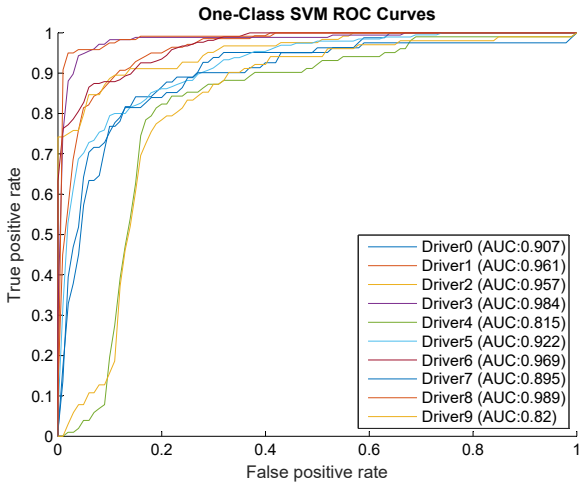

Fig. 4. ROC Curves for One Class SVM Classification of All Study Participants.

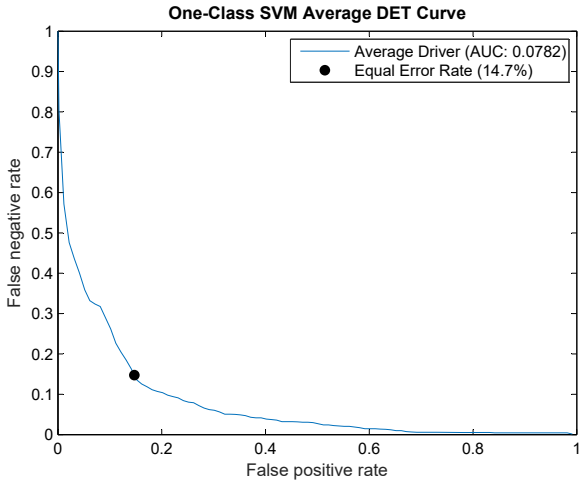

Fig. 5. Average DET Curve for One Class SVM Classification.

a less specific question than the multi-class example. The multiclass model is essentially asking "Are you driver X or driver Y," while the one-class model asks "Are you driver X or a different driver?"

D. Time To Detection

We selected one false positive per driving day as a reasonable performance target. According to a recent study [30], the typical driver spends 46 minutes in his or her vehicle per day on average. Our model works using a 10 second sampling interval. Thus, to achieve a false alarm occurrence rate of one per 46 driving minutes would require a FPR of $\frac{1 \text{ FP}}{46 \text{ minutes}} \times \frac{1 \text{ minute}}{60 \text{ seconds}} \times \frac{10 \text{ seconds}}{1 \text{ sample}} = 0.00362$ or 0.362%. At this very restrictive FP rate, our oc-SVM model of driving behavior is capable of performing driver classification with a detection rate of 19.5%. This means that in any one particular time window there is a 80.5% change of an illicit driver avoiding detection. The following equations calculate how

many 10 second samples would be required to ensure that any unauthorized drivers are detected with at least 95% confidence:

$$\begin{aligned} 0.805^x &< 0.05 \\ x &< \log(0.05)/\log(0.805) \\ x &< 13.81 \end{aligned}$$

From an operational perspective our oc-SVM model of driving behavior can successfully detect illicit vehicle usage with 95% accuracy after 14 samples, or 2 minutes and 20 seconds of driver data collection, while keeping false alerts to once per driving day at most. These results show that identifying drivers is feasible in practice with active behavior modeling without incurring any significant computation or high false positives.

V. Conclusion

To summarize, this paper introduces a novel approach to improve vehicular security in the form of driver modeling for enhanced authentication. We propose to identify drivers via unique characteristics which emerge as they operate a vehicle. We performed a preliminary data collection effort with 10 human subjects in which they completed a simulated driving task while recording their activity. We successfully constructed models of driving activity via extracted features, namely pedal control, steering, speed, and distance traveled. The results of our experiment and modeling effort yield an average EER of 14.7%, implying a time-to-detection of 2 minutes and 20 seconds at 95% confidence with at most one false alert per day of driving.

These results provide strong evidence in support of our hypothesis that drivers can be identified by observing the manner in which a vehicle is operated. As future work, we intend to explore different combinations of features, modeling algorithms, and parameters in order to improve our classification results. Furthermore, given the success of this experiment we intend to pursue a larger scale user study which includes an analysis of different features such as the alertness level of the driver.

Acknowledgment

This research was partially supported by University Transportation Research Center Region II Faculty Initiated Research Program (Grant No. 49198-33-27) and National Science Foundation (NSF) (Grant No. CNS-1559652). The user study protocol was reviewed and approved by the New York Institute of Technology Institutional Review Board.

References

[1] A. Schmitt, "Report: Traffic studies systematically overstate benefits of road projects," http://usa.streetsblog.org/2012/07/06/report-traffic-studies-systematically-overstate-the-benefits-of-road-projects, 2012.

[2] A. for Toll-Free Interstates, "Studying and forecasting tolls is inefficient, unproductive and expensive," http://www.tollfreeinterstates.com/resources, 2016.

[3] M. S. Nicolaisen and P. A. Driscoll, "Ex-post evaluations of demand forecast accuracy: A literature review," *Transport Reviews*, vol. 34, no. 4, 2014.

[4] R. Bishop, *Intelligent vehicle technology and trends*, 2005.

[5] E. C. for Transportation, "Emerging technology trends in transportation," https://www.enotrans.org/wp-content/uploads/EmergingTech.v13.pdf, 2016.

[6] Progressive Corporation, "Snapshot Common Questions," http://www.progressive.com/auto/snapshot-common-questions, 2014.

[7] C. Troncoso, G. Danezis, E. Kosta, J. Balasch, and B. Preneel, "Pripayd: Privacy-friendly pay-as-you-drive insurance," *Dependable and Secure Computing, IEEE Transactions on*, vol. 8, no. 5, pp. 742–755, 2011.

[8] K. Ozbay, "Using big data to identify hotspots of pedestrian crashes in manhattan," in *UTRC Ground Transportation Technology Symposium*, 2014.

[9] L. Hartley, N. R. T. Commission *et al.*, *Review of fatigue detection and prediction technologies*. National Road Transport Commission Melbourne, Australia, 2000.

[10] C. *et al.*, *Sleep Disorders and Sleep Deprivation:: An Unmet Public Health Problem*. National Academies Press, 2006.

[11] M. Jensen, J. Wagner, and K. Alexander, "Analysis of in-vehicle driver behaviour data for improved safety," *International journal of vehicle safety*, vol. 5, no. 3, pp. 197–212, 2011.

[12] M. Salemi, "Authenticating drivers based on driving behavior," Ph.D. dissertation, Rutgers University-Graduate School-New Brunswick, 2015.

[13] C. Miller and C. Valasek, "Adventures in automotive networks and control units," in *DEF CON Hacking Convention*, 2013.

[14] S. TANG *et al.*, "The legal and social implications of real-time drivers physiological state monitoring facilities," *Asian Journal of Management and Humanity Sciences*, vol. 3, pp. 70–79, 2008.

[15] D. Jergler, "Researchers question privacy of usage-based auto insurance, 2013," http://www.insurancejournal.com/news/national/2013/10/02/307073.htm.

[16] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1052–1068, 2004.

[17] C. Karatas, L. Liu, H. Li, J. Liu, Y. Wang, S. Tan, J. Yang, Y. Chen, M. Gruteser, and R. Martin, "Leveraging Wearables for Steering and Driver Tracking," in *Proceedings of IEEE International Conference on Computer Communications (IEEE INFOCOM 2016)*, Apr. 2016.

[18] I. D. Markwood and Y. Liu, "Vehicle self-surveillance: Sensor-enabled automatic driver recognition," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*. ACM, 2016, pp. 425–436.

[19] M. Enev, A. Takakuwa, K. Koscher, and T. Kohno, "Automobile driver fingerprinting," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 1, pp. 34–50, 2016.

[20] C. Troncoso, G. Danezis, E. Kosta, J. Balasch, and B. Preneel, "Pripayd: Privacy-friendly pay-as-you-drive insurance," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 5, pp. 742–755, 2011.

[21] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, T. Kohno *et al.*, "Comprehensive experimental analyses of automotive attack surfaces." in *USENIX Security Symposium*, 2011.

[22] T. Halevi, S. Lin, D. Ma, A. K. Prasad, N. Saxena, J. Voris, and T. Xiang, "Sensing-enabled defenses to rfid unauthorized reading and relay attacks without changing the usage model," in *2012 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2012, pp. 227–234.

[23] T. Halevi, H. Li, D. Ma, N. Saxena, J. Voris, and T. Xiang, "Context-aware defenses to rfid unauthorized reading and relay attacks," *IEEE Transactions on Emerging Topics In Computing*, 2013.

[24] S. Mayhew, "Ford considers biometric authentication for future vehicles," http://www.biometricupdate.com/201409/ford-considers-biometric-authentication-for-future-vehicles.

[25] A. Shamir and N. Van Someren, "Playing hide and seekwith stored keys," in *International conference on financial cryptography*. Springer, 1999, pp. 118–124.

[26] R. E. Helm, "Motor vehicle liability insurance: A brief history," *St. John's Law Review*, vol. 43, no. 1, p. 2, 2012.

[27] OpenDS, "The flexible open source driving simulation," http://opends.de.

[28] M. Welling, "Fisher linear discriminant analysis."

[29] J. Zhao, K. Lu, and X. He, "Locality sensitive semi-supervised feature selection," *Neurocomputing*, vol. 71, no. 10, pp. 1842–1849, 2008.

[30] J. Hall, "New study reveals when, where and how much motorists drive," http://newsroom.aaa.com/2015/04/new-study-reveals-much-motorists-drive/, 2015.