In-Car Driver Authentication Using Wireless Sensing

Sai Deepika Regani Qinyi Xu Beibei Wang Min Wu K. J. Ray Liu
Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA

***Abstract*— Automobiles have become an essential part of everyday lives. In this work, we attempt to make them smarter by introducing the idea of in-car driver authentication using wireless sensing. Our aim is to develop a model which can recognize drivers automatically. Firstly, we address the problem of "changing in-car environments", where the existing wireless sensing based human identification system fails. To this end, we build the first in-car driver radio biometric dataset to understand the effect of changing environments on human radio biometrics. This dataset consists of radio biometrics of five people collected over a period of two months. We leverage this dataset-to create machine learning (ML) models that make the proposed system adaptive to new in-car environments. We obtained a maximum accuracy of 99.3% in classifying two drivers and 90.66% accuracy in validating a single driver.**

***Index Terms*— Driver authentication, human radio biometrics, wireless sensing, radio shot, human identification.**

I. INTRODUCTION

By deploying tremendous connected smart devices and analyzing the gathered data, the Internet of Things enables evolutionary changes in every aspect of people's daily life, including the emerging smart automobiles. One of the important and interesting aspects of smart automobiles is driver authentication which enables automatic adjustment of internal settings in automobiles such as seat and mirror positions, temperature etc., that are specific to an individual and can be operated without the need for a key.

Traditional approaches such as fingerprint matching, face recognition, iris technology and many more [1], mostly utilize techniques of image processing or computer vision to identify people. Human identification has also been done by observing the gait of a person, which is inapplicable in cars [2]. All these techniques require video or images taken from a camera to perform human identification and have the drawback of potential privacy leakage.

On the other hand, due to the ubiquitous deployment of wireless technology, wireless sensing becomes an innovative solution to many IoT applications [3], including smart car systems.

Previous works used wireless sensing for driver activity recognition [4], gait recognition [5], [6] etc. However, to the best of our knowledge, this is the first attempt to use wireless sensing for in-car driver authentication.

While gait can be modified and cannot be assured to be the same every time, radio biometrics of a person is more reliable. The first work of this kind [7], used time reversal resonating strength (TRRS) to compare different radio biometrics embedded and recorded in the wireless channel state information (CSI). However, it assumes constant environment which is not the case in real world scenarios. The CSI is sensitive to the environment and the problem of changing environment poses a great challenge to the method. In the proposed work, we investigate and attempt to solve this problem by leveraging radio signature collections of drivers for different in-car environments. This led to the development of the first in-car driver radio biometric dataset.

The key contributions of the paper are summarized below.

- We propose the first in-car driver authentication system using WiFi.
- We build the first driver radio biometric dataset consisting of radio signatures of five people collected over a period of two months.
- Using the above dataset, we develop machine learning (ML) models which can adapt to in-car environmental changes and improve the accuracy of driver authentication.

Overall, this paper is organized as follows. Section II discusses the main challenges. Section III describes the in-car driver authentication system and the dataset preparation. Section IV summarizes the evaluation methodologies. Finally, Section V presents the performance of different ML models for two driver authentication and single driver validation.

II. CHALLENGES

The procedure of recording a radio signature is called radio shot [7]. The similarity of two CSIs can be defined by the TRRS. For two Channel Frequency Responses (CFRs) $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$, the TRRS in the frequency domain is given by [7]:

$$TRRS(\boldsymbol{h_1}, \boldsymbol{h_2}) = \frac{max_{\phi}|\sum_{k=0}^{L-1} h_1[k]h_2[k]^* e^{jk\phi}|^2}{(\sum_{l=0}^{L-1}|h_1[l]|^2)(\sum_{l=0}^{L-1}|h_2[l]|^2)} \quad (1)$$

where L is the number of sub-carriers and the $()^*$ operator denotes the conjugate operator. The higher the TRRS is, the more similar the two CFRs are, and thus the more similar the two radio biometric samples are. The TRRS based approach [7], proposed and proved the existence of human radio biometrics and assumed that the indoor environment remains the same throughout the experiment.

However, in a practical case, the indoor environment changes rapidly, leading to the varying CSI and a performance degradation due to the mismatch of CSI. This is because, the human radio biometrics are embedded in the wireless CSI which is highly correlated with the propagation environment. To study the changing environments in a car, we have recorded the CSI of an empty car for three months. Fig. 1 shows the CSI variation of an empty car in terms of TRRS value, calculated with the reference of day one CSI.

In general, the TRRS value decreases along time, as environmental changes accumulate, leading to an environment that differs from the original one.

Fig. 2, shows the TRRS matrix between CSI collected during the radio shots of two people on two different days, A and B. The similarity of any two CSIs can be obtained from the corresponding value in the TRRS matrix. The similarity between the two empty in-car environment CSIs dropped to 0.69. Using the existing TRRS matching technique [7], human 2 on day B would either not be identified if there is a threshold on the similarity or would be recognized as human 1 since the TRRS between the two radio shots is higher(i.e., 0.73) than that of between the same person (i.e., 0.57). In this case, the changed environment caused a lower TRRS which lead to a mismatch. Hence, there is a need for a long term test and a broader study to understand the behaviour of human radio biometrics under changing environments.

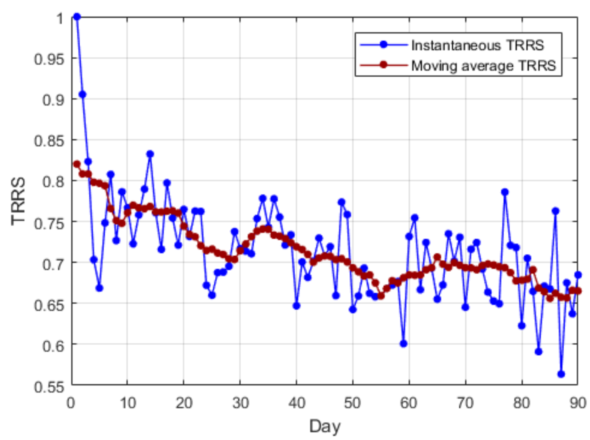

Fig. 1. Degree of environment change inside the car

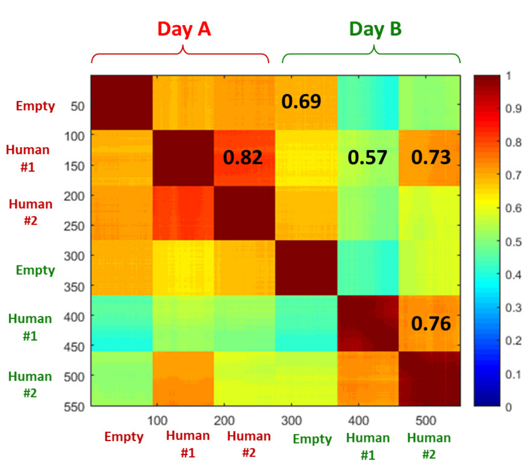

Fig. 2. TRRS matrix for radio shots on different days for two people

III. IN-CAR DRIVER AUTHENTICATION

In this work, we build a human radio biometric dataset, by collecting radio shots of five people over two months. On each day, for each test subject, four radio shots are taken in the morning and evening each, in a car parked at different locations in a public parking lot. By doing so, a total of 60 different environments have been considered. The location of the transceivers and the driver are as shown in Fig. 3. Our goal is to obtain the identity of the driver on a new day, given the radio shots from the past days/environments.

Fig. 3. Location of transceivers in the car

To evaluate the classification performance of the machine learning techniques over a limited dataset, cross-validation techniques [8] are used. In this technique, the entire dataset is divided into K parts and K experiments are performed with each part as testing data and the remaining $K-1$ parts as the training data, as explained in Fig. 4. When the amount of data is limited, any particular split of train and test data, may not include all the possible variations. In such cases, the accuracy values obtained would be biased and dependent on the choice of division. K-fold validation allows us to create a more unbiased estimate of the accuracy. In this paper, we denote the number of folds as K_v.

We build the prototype of the proposed driver authentication system on off-the-shelf WiFi chips. We obtained the CSIs using a 3x3 MIMO system under a sounding rate of 30 Hz. The data is collected over a bandwidth of 40 MHz in the 5.2 GHz band with 114 accessible subcarriers. During each radio shot, 90 CSIs are collected, and they are highly correlated and can be used to remove outliers. The obtained CSI matrix per radio shot, is a complex valued matrix of dimensions 3x3x114x90.

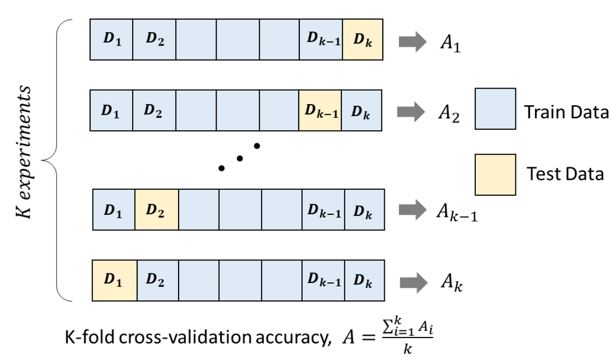

Fig. 4. k-fold validation technique

Data Preprocessing: The CSIs obtained from the radio shot need to be preprocessed to eliminate the phase distortions. We compensate for the linear and the initial phase offsets as discussed in [7]. After the phase alignment, each CSI as a 3x3x114 dimensional complex valued vector is translated to a 2052 dimensional real valued vector. With such a high dimension of feature, the number of parameters is large and machine learning techniques usually require more data to learn. Hence, we perform dimensionality reduction using Principle Component Analysis. Considering about 99% of the variance in the data, the number of dimensions can be reduced to 90.

IV. Evaluation Methodology

A. K- Nearest Neighbour (KNN)

Every new in-car environment presents a new instance of the data to the driver authentication system. This points us to the class of instance-based learning methods [9], of which KNN is the most popular and the simplest one and we use it as a baseline. For every new radio shot, we find the closest K points from the database and assign the majority identity to the test sample. We select the value of K based on the maximum average K_v-fold accuracy. Fig. 5 shows the mean and standard deviation of the accuracy from K_v experiments. Based on the results, we choose the value $K = 3$ as it has the maximum accuracy and lowest standard deviation for this particular classification.

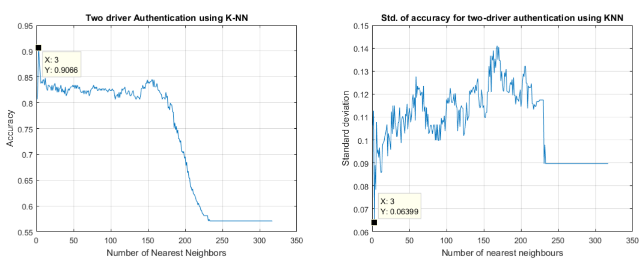

Fig. 5. Mean and standard deviation of accuracies for different number of nearest neighbors. The optimal value of K is 3 for this classification.

B. Support Vector Machine (SVM)

SVM technique tries to find a separating hyper plane between two classes that maximizes the margin between the plane and the closest point from either class. If the data is not linearly separable, we can use a “kernel-trick” to project the data into a very high dimensional plane using non-linear kernels [10]. In the proposed in-car driver authentication system, we study the performance of SVM using both linear and radial basis function (RBF) kernels.

C. Neural Network (NN)

We need a system which is adaptive and can learn the human radio biometrics under different environments. Deep neural networks have been proved to do these tasks really well in the computer vision field [11]. In the proposed in-car driver authentication system, we adopt the NN technique with ReLU activation function. The output of the network is the class probabilities which can be used to obtain the identity of the person. We first performed dimensionality reduction over the collected radio biometrics and it allows us to use a smaller network with fewer parameters. The network architecture is shown in Fig. 6.

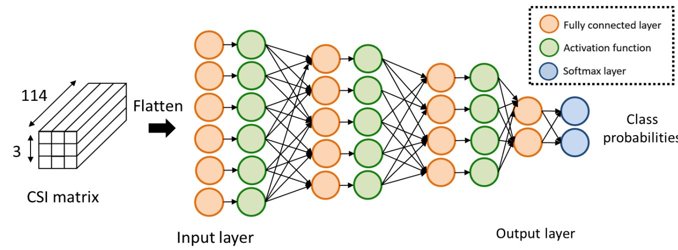

Fig. 6. Neural network architecture with two fully connected hidden layers and ReLU activation

D. Grouping

During the process of capturing the radio shot, varied seating positions of test subjects may cause minor changes in the radio signatures. This can be clearly seen even for radio shots which are taken almost at the same time i.e., with negligible change in the in-car environment. To make the system robust to such variations, we take multiple radio shots for each test subject every time, during training and testing phases. In the testing phase, we predict the class using all the realizations and use the combined class probabilities to obtain the driver identity. Fig.7 explains the grouping technique in case of a neural network. For each test subject, we collect 4 radio shots and index them as $i, i = 1, 2, 3, 4$. Let P_{Ai} and P_{Bi} represent the predicted class probability of the i^{th} radio shot under class A and class B, respectively. Then the identity of the test subject is determined as class A if $\Sigma P_{Ai} > \Sigma P_{Bi}$ and vice versa. The accuracy values obtained using grouping, indicate the maximum accuracy that can be achieved on a new day under a new environment.

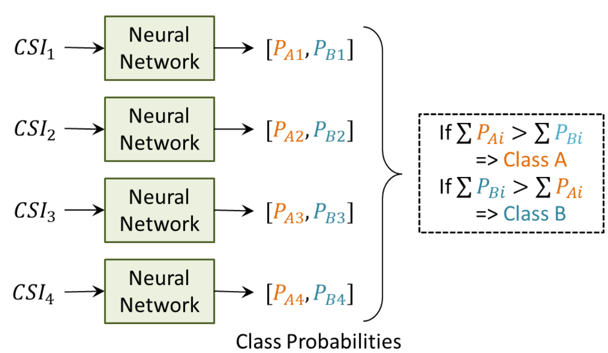

Fig. 7. Group decision for a neural network

V. Experimental results and Discussions

To evaluate the performance of the proposed in-car driver authentication system, we consider two scenarios of driver authentication.

A. Two Driver Authentication

In this scenario, the proposed systems learn the radio biometrics of two drivers and tries to differentiate between them. As a binary classification problem, it allows the system to automatically recognize one of the two registered drivers of a car and apply driver specific adjustments of the seat positions, mirrors and temperature. The car can also be used without a key and the driver authentication system can serve as a security feature.

Table I and Table II shows the K_v-fold accuracy values, obtained using different ML techniques and grouping, for all combinations of two drivers from the five-driver database. We observe that in most cases, the NN approach learns better. Also, the grouping technique improves the performance of all the techniques. The best pair gives an accuracy of 99.36% and 94.88%, with and without grouping respectively.

TABLE I
PERFORMANCE ON TWO DRIVER AUTHENTICATION

Classes	K-NN	Linear SVM	SVM-RBF	NN
A-B	86.29	91.67	89.58	89.58
A-C	90.30	92.08	93.39	**94.88**
A-D	87.66	94.06	92.19	92.79
A-E	88.82	91.87	94.69	94.23
B-C	85.20	86.93	86.40	90.44
B-D	76.70	86.61	86.93	86.60
B-E	85.5	88.85	91.35	91.88
C-D	81.80	86.87	88.80	91.24
C-E	69.08	75.31	74.48	80.40
D-E	80.15	89.22	89.17	91.50

TABLE II
PERFORMANCE ON TWO DRIVER AUTHENTICATION WITH GROUPING

Classes	KNN	SVM-Linear	SVM-RBF	NN
A-B	86.40	93.75	91.45	96.58
A-C	91.00	94.17	94.17	**99.36**
A-D	88.81	94.80	93.54	98.08
A-E	89.25	91.46	95.21	**99.36**
B-C	87.06	89.37	87.91	96.37
B-D	74.12	87.30	89.37	95.51
B-E	85.53	90.62	93.33	94.88
C-D	83.77	85.00	89.17	95.10
C-E	65.57	75.42	73.33	84.19
D-E	80.70	88.75	91.04	96.80

B. Single Driver Validation

In this scenario, we recognize one registered driver of the car. This is a one-vs-many classification problem. During the training phase, we train the system to differentiate between person A and person B, C, D. During the testing phase, we test if the system can differentiate person A and persons E, F. The classification acccuracy obtained using RBF-SVM is 90.66%.

C. Similarity of radio shots

Two people who have higher similarity in the radio signatures in the same environment, have a lower classification accuracy. For example, in Table III, we show the TRRS between the radio shots of two individuals averaged over all environments. A lower inter-class TRRS corresponds to a higher classification accuracy.

TABLE III
CONSISTENCY FACTOR

Classes	Average TRRS on same day	Accuracy%
A-E	0.7094	99.36
C-E	0.7773	84.19

D. Smart system: Learning with time

The proposed system is 'smart', in the sense that, it learns more and more with time i.e., more data. Initially, when the data is very limited, instance based methods or SVM can be used to do the classification. When a good amount of data accumulates, we train a NN to perform driver authentication. In Fig. 8, we show a moving average of the performance of a NN with an increasing amount of data. The accuracy values do not show a sturdy increase because the feature contains casual time-varying pattern that depends on the new empty car environment and will only be present in the test case.

Overall, we speculate that the ML models learn more environment independent and human specific features with time. This became possible by training the model using radio biometrics collected from a large number of different environments present in the driver radio biometric dataset.

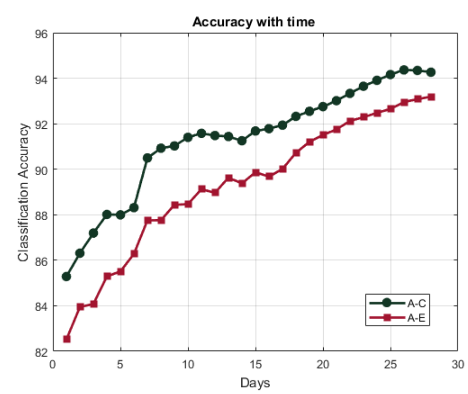

Fig. 8. Smart learning: The performance of the driver authentication system increases as it learns more with time.

VI. CONCLUSIONS

In this work, we have proposed the first in-car driver authentication system with commercial WiFi devices. The proposed system gave an accuracy of 90.66% and 99.36% for single driver validation and two driver authentication, respectively. Moreover, this is the first work to study the behaviour of human radio biometrics in a long term. We believe this research would open up new applications of wireless sensing in security systems and biometrics.

References

[1] D Bhattacharyya, R Ranjan, F Alisherov, M Choi, et al. Biometric authentication: A review. *International Journal of u-and e-Service, Science and Technology*, 2(3):13–28, 2009.

[2] A. Kale, A. Sundaresan, A. N. Rajagopalan, N. P. Cuntoor, A. K. Roy-Chowdhury, V. Kruger, and R. Chellappa. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9):1163–1173, Sept 2004.

[3] B. Wang, Q. Xu, C. Chen, F. Zhang, and K. J. R. Liu. The promise of radio analytics: A future paradigm of wireless positioning, tracking, and sensing. *IEEE Signal Processing Magazine*, 35(3):59–80, May 2018.

[4] S Duan, T Yu, and J He. Widriver: Driver activity recognition system based on wifi csi. *International Journal of Wireless Information Networks*, 25(2):146–156, Jun 2018.

[5] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere. Wifi-id: Human identification using wifi signal. In *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 75–82, May 2016.

[6] W Wang, A X. Liu, and M Shahzad. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, pages 363–373, New York, NY, USA, 2016. ACM.

[7] Q. Xu, Y. Chen, B. Wang, and K. J. R. Liu. Radio biometrics: Human recognition through a wall. *IEEE Transactions on Information Forensics and Security*, 12(5):1141–1155, May 2017.

[8] M Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological)*, pages 111–147, 1974.

[9] D W. Aha, D Kibler, and M K. Albert. Instance-based learning algorithms. *Machine Learning*, 6(1):37–66, Jan 1991.

[10] I Steinwart and A Christmann. *Support vector machines*. Springer Science & Business Media, 2008.

[11] M Long, Y Cao, J Wang, and M I Jordan. Learning transferable features with deep adaptation networks. *arXiv preprint arXiv:1502.02791*, 2015.