



IDENTIFICATION OF DRIVING CHARACTERISTICS WITH MARKOV PROCESS MODEL

Tuba Nur SERTTAŞ^{1,*}, Ömer Nezh GEREK², Fatih Onur HOCAOĞLU³

¹ Electrical Electronics Engineering, Technology Faculty, Afyon Kocatepe University, Afyonkarahisar, Turkey

² Electrical Electronics Engineering, Engineering Faculty, Eskişehir Technical University, Eskişehir, Turkey

³ Electrical Engineering, Engineering Faculty, Afyon Kocatepe University, Afyonkarahisar, Turkey

ABSTRACT

Considering the role of drivers in traffic, it is an important advantage to know the driver characteristics in advance. Taking this characteristic into consideration, the driver can be warned socially and economically. In this study, taking into account this situation, driving characteristics are removed within certain drivers and the drivers are divided into two classes as calm and aggressive. The data recorded via the smartphone is used directly when classification. By applying Markov process method to the data, the drives are classified with 73% accuracy.

Keywords: Driver classification, Markov Process, Drive data analysis

1. INTRODUCTION

Increasing population and developing technology brought with it inevitable end of traffic and vehicle growth has emerged with many issues that need to be considered and examined. Thanks to their comfort, drivers can travel longer and consequently distractions can occur while driving. The increasing number of vehicles leads to more resource consumption and a more polluted world. Stress factors during the day make drivers more nervous and impatient in traffic. With such traffic results, accident rates, gas amounts in the air we breathe, national resources spent increase. At this stage, research has been made on the drivers by using the opportunities offered by the technology and it has been tried to cope with the mentioned negativities. In the literature, the issue of distraction, the issue of stress, the type of driver and energy saving have been dealt with many times and from different perspectives.

This study presents a novel approach to classify drivers according to their driver behaviors. Such a classification is beneficial for developing energy management systems, searching the underlying process of traffic accidents and building intelligent insurance tariff mechanisms. There are many methods used in the classification of driver characteristics. In this study, Markov Process method is used as classification method. While the use of Hidden Markov Method (HMM) is generally preferred in the literature, Markov Process has been preferred because it can directly use raw data as a case in order to minimize complexity and these situations are directly related to the results.

Level 3 autonomous driving seems to be more viable in the near future [1]. With the new specifications, driver is not required to monitor the vehicle all the time. A PID controller for SBW systems and a robust observer considering Coulomb friction are occurred and analyzed to ensure that the system is asymptotically stable and independent of driver torque parameter uncertainties.

Abnormal driving behaviors are very problematic in the traffic [2]. A driving data recognition method based on a long short-term memory network and convolutional neural network (LSTM-CNN) is proposed. The extreme points in the real driving data are detected and statistically analyzed. In another study [3], acceleration, deceleration, steering and lane changing data of drivers are evaluated in statistical parameters. Successful evaluation of lane changing maneuvers may provide early warning

*Corresponding Author: tngul@aku.edu.tr

Received: 01.07.2019 Published: 15.06.2020

and help drivers in critical decision making. Deep learning techniques are also preferred in the recent driving behavior studies [4]. They are proved to outperform conventional machine learning classifiers. The empirical trajectory dataset is utilized to perform the models.

Support vector machines have accurate results in the driving classification [5]. Radial Basis Function kernel is proposed monitoring a highly sensitive frame for a large number of drivers. Classification accuracy may increase up to 90% between 10-24 drives with SVM method. GPS based driving experience data is utilized in many applications [5], [6]. Wireless sensing to collect biometrics of drivers to classify driver behavior is currently studied and machine learning methods are performed [7]. Frequency response function (FRF) estimates reveal how drivers use visual preview, lateral position feedback, and heading feedback for control [8].

Hidden markov models are used in a wide range of areas from audio processing to signal and image processing, and are thus applied to operations such as driver recognition, routing, and driver behavior identification. HMMs that are already used in speech processing find application in various pattern recognition areas such as learning and recognizing space for mobile robots [9], signature recognition [10] and learning human behavior for teleoperations of robots [11]. Human behavior is both stochastic and dynamic. Dynamic models using Kalman filter are used to reveal these two characteristics of human behavior. This model feeds prediction errors in HMMs and predicts vehicle output from drivers input data. [12] showed that only the first 2 seconds of the driving data set would be sufficient to predict a driver's behavior. This proves how useful the use of HMMs is in such studies. In a study on driving event recognition, the data obtained from the vehicle were used directly. Among these data, event recognition was provided with 93.8% accuracy using only speed and acceleration data [13]. In addition to the data used in this study, steering wheel angle is known for the sub-maneuvers described as 'drivemes' in a study using headway data. In addition, this study showed that HMM can make a very good definition even at complex intersections. In [15], where it is thought that the interaction of vehicles in traffic is an important factor, the measurements of the vehicle used and the surrounding vehicles were measured on a crossroad and proved that only one HMM is sufficient for identification. This study presents an opposing approach to modeling driver behavior. This approach starts from the top with a single HMM and separates the initial HMM structure into fragmenting and clustering understanding structures. In this study, three main clusters have been identified and a new approach for defining driver behavior for intersections is presented. In [16], the additional transition possibilities defined between the good and bad performance sequences are used. Thus, both recognition and performance evaluation were proposed at the same time. The mentioned studies show how useful HMM is in driver behavior modeling and provide a supportive attitude for the studies to be performed. HMM is considered as a suitable method for driving behavior modeling for the following two reasons.

_HMMs can model the stochastic structure of driving behavior by providing statistical corrections when proposing temporary modeling

_Variations in the driving signals can be modeled or suppressed to meet the requirements of the desired task.

In [17], using only three CAN-Bus signals (i.e., vehicle speed, steering wheel angle and brake force) three different maneuvers (left turn, right turn and lane change) are recognized. Hidden Markov Models are employed in bottom-to-top (BtT) and top-to-bottom (TtB) approaches in a complimentary way; latter being used to discover the optimum number of states for each maneuver verifying the intuitive BtT method. Hidden Markov models (HMM) are used to identify a vehicle's maneuver sequence.

In this study, in the second chapter, the classification method used, Markov model and the data used are mentioned. 39 different drives are carried out on a predetermined driving route. During these rides, data is recorded with the smartphone application. In the third chapter, the drive characteristic values obtained from the classification are given and the classification results are presented. The drivers are divided into two basic groups as calm and aggressive. Using driving data, the states for the Markov process method are created and as a result the drives are classified with 73% accuracy.

2. MATERIALS AND METHODS

2.1. Markov Process

The Markov chain is a stochastic process that has a memory-less feature, unlike other classical processes. This means that the future state does not depend on other states except for the present state.

The Markov Property: For any positive integer n and possible states $i_0, i_1, i_2, \dots, i_n$ of the random variables.

$$P(X_n = i_n | X_{n-1} = i_{n-1}) = P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \quad (1)$$

That is, there is no need for information other than the information of the previous situation to determine the probability distribution of the present situation. This definition is broader than the one explored above, as it allows for non-stationary transition probabilities and therefore time-inhomogeneous Markov chains; that is, as time goes on (steps increase), the probability of moving from one state to another may change.

Transition Matrices: A Transition matrix P_t for Markov chain $\{X\}$ at time t is a matrix containing knowledge on the probability of transitioning between states. In particular, given an ordering of a matrix's rows and columns by the state space S , the $(i, j)^{th}$ element of the matrix P_t is given by

$$P(t)_{i,j} = P(X_{t+1} = j | X_t = i) \quad (2)$$

This means each row of the matrix is a probability vector, and the sum of its entries is 1. Transition matrices have the property that the product of subsequent ones describes a transition along the time interval spanned by the transition matrices. That is to say $P_0 \cdot P_1$ has in its $(i, j)^{th}$ position the probability that $X_2 = j$ given that $X_0 = i$. And, in general, the $(i, j)^{th}$ position of $P_t \cdot P_{t+1}, \dots, P_{t+k}$ is the probability $P(X_{t+k+1} = j | X_t = i)$.

The k -step transition matrix is $P_t^{(k)} = P_t \cdot P_{t+1}, \dots, P_{t+k-1}$ and, by the above, satisfies

$$P_t^{(k)} = \begin{pmatrix} P(X_{t+k} = 1 | X_t = 1) & P(X_{t+k} = 2 | X_t = 1) & \dots & P(X_{t+k} = n | X_t = 1) \\ P(X_{t+k} = 1 | X_t = 2) & P(X_{t+k} = 2 | X_t = 2) & \dots & P(X_{t+k} = n | X_t = 2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ P(X_{t+k} = 1 | X_t = n) & P(X_{t+k} = 2 | X_t = n) & \dots & P(X_{t+k} = n | X_t = n) \end{pmatrix} \quad (3)$$

2.2. Driver Classification

In order to ensure that all driving conditions encountered in daily life of the driving area, conditions such as pedestrian crossing and speed limit are taken into consideration. The driving area provides all traffic maneuvers such as U-turn, lane change, bump and pit crossing. The data of the drivers are recorded with the smart phone application in 39 drives. Speed, 3 axis angular velocity, 3 axis acceleration etc. are obtained from drives. Since this study aimed to classify with a small amount of data, not all data types are used. The speed and angular velocity values are selected from the data to classify. X-axis angular velocity represents the angle of rotation of the front panel of the vehicle. Therefore, the angular velocity data used indicates how hard the driver performs the turns and lane changes.

In the Markov model, which is considered as classification method, 20 states listed in Table 1 have been determined considering the distributions of speed and angular velocity values of drivers. Some of the conditions in the table have never been used during the testing phase. However, the status ranges are kept low so that more accurate case analysis can be performed from the available data.

The states that arise when all drivers complete the specified route are given in Figure 1. When the graphs in the figure are examined, the situations used by the two groups do not show much difference. However, it is seen that situations that are never used for calm drivers are used during the test drive of aggressive drivers. While these graphs cannot be used directly during the classification of the drivers, the probability transition matrices based on these graphs contain important information in terms of classification. The difference in number is due to the completion of the test course at different times. A calm driver completes the same course in approximately 6 minutes, while the aggressive driver completes in 4 minutes.

Table 1. Classification of Driver

States	Derivation of speed(ΔV)	Angular velocity(ω_x)
State 1	$\Delta V < 0$	$\omega_x < -2$
State 2	$\Delta V < 0$	$-2 \leq \omega_x < -1.5$
State 3	$\Delta V < 0$	$-1.5 \leq \omega_x < -1$
State 4	$\Delta V < 0$	$-1 \leq \omega_x < -0.5$
State 5	$\Delta V < 0$	$-0.5 \leq \omega_x < 0$
State 6	$\Delta V < 0$	$0 \leq \omega_x < 0.5$
State 7	$\Delta V < 0$	$0.5 \leq \omega_x < 1$
State 8	$\Delta V < 0$	$1 \leq \omega_x < 1.5$
State 9	$\Delta V < 0$	$1.5 \leq \omega_x < 2$
State 10	$\Delta V < 0$	$2 \leq \omega_x$
State 11	$\Delta V \geq 0$	$\omega_x < -2$
State 12	$\Delta V \geq 0$	$-2 \leq \omega_x < -1.5$
State 13	$\Delta V \geq 0$	$-1.5 \leq \omega_x < -1$
State 14	$\Delta V \geq 0$	$-1 \leq \omega_x < -0.5$
State 15	$\Delta V \geq 0$	$-0.5 \leq \omega_x < 0$
State 16	$\Delta V \geq 0$	$0 \leq \omega_x < 0.5$
State 17	$\Delta V \geq 0$	$0.5 \leq \omega_x < 1$
State 18	$\Delta V \geq 0$	$1 \leq \omega_x < 1.5$
State 19	$\Delta V \geq 0$	$1.5 \leq \omega_x < 2$
State 20	$\Delta V \geq 0$	$2 \leq \omega_x$

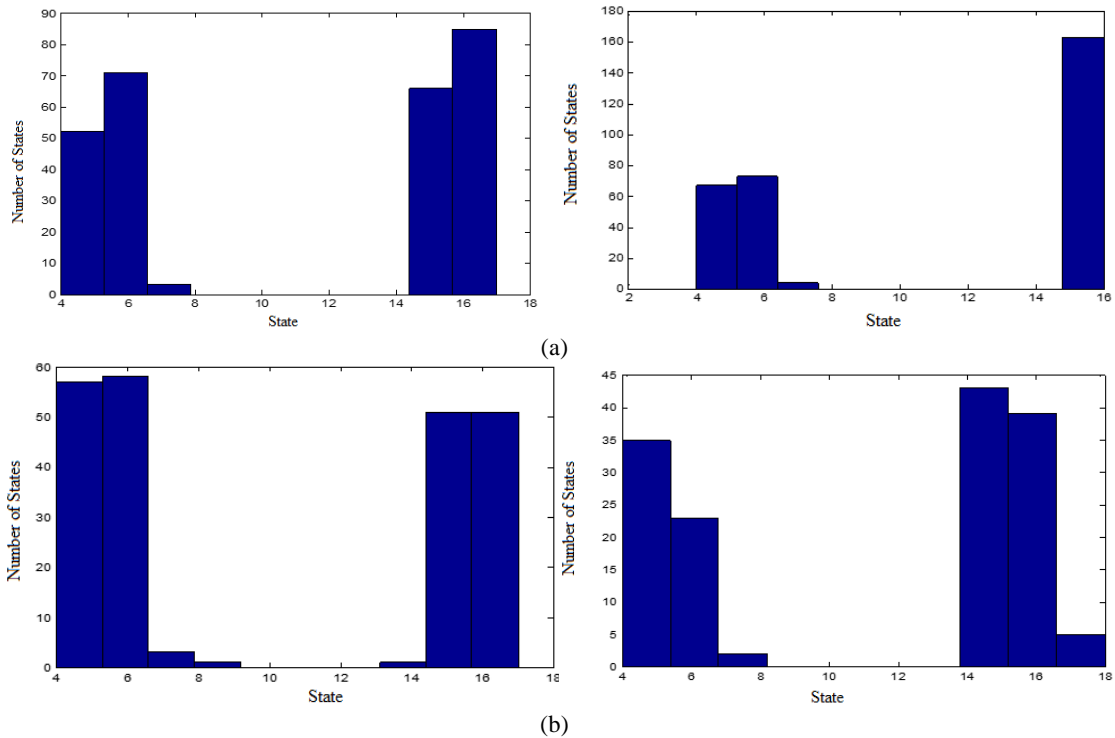


Figure 1. The states of different drivers(a-calm, b-aggressive)

Transition probability matrices are obtained for all drives in accordance with the above state information. Sample matrices are for two drivers belongs to different classes given in Table 2 and Table 3. Unused cases are not included in the table.

Table 2. An example transition probability matrix of a calm driver

State No	4	5	6	7	8	14	15	16	17	18	19	20
4	0	0	0	0	0	0	1	0	0	0	0	0
5	0,04	0,48	0,26	0	0	0	0,12	0,1	0	0	0	0
6	0	0,15	0,63	0,01	0	0	0,07	0,12	0	0	0	0
7	0	0	0	0,33	0	0	0	0	0,66	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0,18	0,03	0	0	0	0,54	0,24	0	0	0	0
16	0	0,04	0,13	0,01	0	0	0,2	0,60	0	0	0	0
17	0	0	0	0	0	0	0	1	0	0	0	0

Table 3. An example transition probability matrix of a aggressive driver

State No	4	5	6	7	8	14	15	16	17	18	19	20
4	0,42	0,28	0	0	0	0,14	0,14	0	0	0	0	0
5	0,06	0,26	0,32	0	0	0	0,22	0,14	0	0	0	0
6	0	0,27	0,34	0	0	0	0,10	0,24	0,01	0	0	0
7	0	0	0	0,33	0,33	0	0	0	0,33	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	1	0	0	0	0	0
15	0,02	0,29	0,13	0	0	0	0,37	0,17	0	0	0	0
16	0	0,08	0,31	0,04	0	0	0,22	0,33	0	0	0	0
17	0	0	0	0	0	0	0,66	0,33	0	0	0	0

21 of the test drivers exposed calm driving characteristics and 18 drivers exposed aggressive driving characteristics. The probability transition matrices of these drives show an increase or decrease in approximately the same cells. If the amount of increase and decrease shows differences, the tendency to change is parallel.

3. RESULTS

The probability represents the tendency of the transition matrix to maintain the same state of the diagonal components. It is therefore considered to be the best expression of calm driver behavior. It means the transition of the most distant states with the corner elements of the matrix. This value is expected to be greater for an aggressive driver. Therefore, the probability transition matrix is multiplied by a determined coefficient matrix and its weighted arithmetic mean is taken.

Weighted Arithmetic Mean: The weighted arithmetic mean of a set of numbers X_1, X_2, \dots, X_N with respective weights of w_1, w_2, \dots, w_N is defined as

$$\bar{X} = \frac{w_1X_1+w_2X_2+\dots+w_NX_N}{w_1+w_2+\dots+w_N} \tag{4}$$

Table 4 shows the weighted arithmetic mean values of all driving.

Table 4. Weighted arithmetic mean of drives

Number of Drive	Mean Value
1 st Drive	34,695
2 nd Drive	33,684
3 th Drive	27,708
4 th Drive	32,711
5 th Drive	41,012
6 th Drive	50,651
7 th Drive	25,397
8 th Drive	35,212
9 th Drive	40,553
10 th Drive	34,930
11 th Drive	50,220
12 th Drive	44,747
13 th Drive	50,179
14 th Drive	52,492
15 th Drive	46,016
16 th Drive	48,204
17 th Drive	51,570
18 th Drive	34,422
19 th Drive	33,180
20 th Drive	24,910
21 th Drive	38,853
22 th Drive	26,992
23 th Drive	27,029
24 th Drive	39,844
25 th Drive	41,168
26 th Drive	46,749
27 th Drive	46,749
28 th Drive	28,588
29 th Drive	35,909
30 th Drive	43,237
31 th Drive	28,343
32 th Drive	49,593
33 th Drive	37,606
34 th Drive	39,026
35 th Drive	38,080
36 th Drive	38,080
37 th Drive	42,738
38 th Drive	25,534
39 th Drive	37,171

A discrimination value is determined according to the distribution of the driver's mean. As a result, 11 out of 39 driving are classified as faulty. The drives are classified with 73% accuracy. While calculating this accuracy percentage, the values in Table 4 are taken into consideration. By determining the weighted average value range of each class, the compatibility of the values in the table with these values is examined. These values are determined by the class assigned to the drivers by the controlling participant during the tests. For example, if the average driving value of a driver in the normal class is 39, this driver is assigned to the wrong class. Because this value is a must for a driver in the aggressive class. Therefore, while addressing this as a biased result, those in the correct range were determined as the correct result. The percentage of accuracy was calculated using following equation.

$$Accuracy = \frac{\text{Number of drivers with weighted driving average in the correct range}}{\text{Total number of drivers}} \quad (5)$$

Accuracy percentages vary between 60.63 and 100 for the classification and identification of drives in the driver classification studies carried out with different methods [2-6,18,19]. Although the

classification percentage is not very high in this study, it is advantageous and different to use the data as raw data without any processing. The data of the driver was used as the data set of the Markov method without any maneuvering or feature extraction. This situation both simplified the application and significantly reduced the number of transactions.

4. CONCLUSION

In this study, a driver classification procedure is carried out in which 39 driving with age and sex different drivers are handled. Drivers are prevented from changing their driving characteristics by not being informed about the content of the study and the driving area. The drivers are divided into two classes by the expert who participated in all the rides. The data recorded with the smartphone application while driving is used as data in the Markov Process method and the classification process is completed. Accuracy analysis is performed by comparing this classification with the classification performed by the expert. According to this result, the drives are classified with 73% accuracy. It has been demonstrated that the classification by using driving data directly will provide high accuracy with the Markov Method instead of the commonly used HMM in the literature. In the future, studies can be made on topics such as vehicle design based on driver classes, energy management system or vehicle insurance fee determination using this classification result.

ACKNOWLEDGEMENTS

This study is supported by Eskişehir Technical University Scientific Research Project numbered 1701F020.

REFERENCES

- [1] Huang C, Li L, Liu Y, Xiao L. Robust Observer Based Intermittent Forces Estimation for Driver Intervention Identification. *IEEE Trans. Veh. Technol.*, Apr. 2020; 69(4): 3628–3640.
- [2] Jia S, Hui F, Li S, Zhao X, Khattak AJ. Long short-term memory and convolutional neural network for abnormal driving behaviour recognition. *IET Intell. Transp. Syst.*, May 2020; 14(5): 306–312.
- [3] Lin X, Zhang K, Cao W, Zhang L. Driver Evaluation and Identification Based on Driving Behavior Data. In *Proceedings - 2018 5th International Conference on Information Science and Control Engineering*, 2019; 718–722.
- [4] Liu X, Liang J, Xu B. A Deep Learning Method for Lane Changing Situation Assessment and Decision Making. *IEEE Access*, 2019; 7:133749–133759.
- [5] Rahim MA, Liu J, Zhang Z, Zhu L, Li X, Khan S. Who is driving? Event-driven Driver Identification and Impostor Detection through Support Vector Machine. *IEEE Sens. J.*, Jun. 2020; 20(12):1.
- [6] Rahim MA, Zhu L, Li X, Liu J, Zhang Z, Qin Z, Khan S, Gai K. Zero-to-Stable Driver Identification: A Non-Intrusive and Scalable Driver Identification Scheme. *IEEE Trans. Veh. Technol.*, Jan. 2020; 69(1): 163–171.
- [7] Regani SD, Xu Q, Wang B, Wu M, Ray Liu KJ. Driver Authentication for Smart Car Using Wireless Sensing. *IEEE Internet Things J.*, Mar. 2020; 7(3): 2235–2246.
- [8] Van der El K, Pool DM, Van Paassen MRM, Mulder M. A Unifying Theory of Driver Perception and Steering Control on Straight and Winding Roads. *IEEE Trans. Human-Machine Syst.*, Apr. 2020; 50(2): 165–175.

- [9] Aycard O, Charpillet F, Foht D and Mari JF. Place learning and recognition using hidden Markov models. In Proc, IEEE Int. Robots Syst., 1997, Grenoble, France, pp. 1741-1746.
- [10] Yang L, Widjaja BK and Prasad R. Application of hidden Markov models for signature verification. Pattern Recognition, 1995; 28:161-169.
- [11] Yang Y Xu Y. Human Action Learning via Hidden Markov Model. IEEE Trans. Syst. Man and Cyber.-Part A: Syst. And Humans, 1997; 27:1.
- [12] Pentland A, Liu A. Modeling and Prediction of Human Behavior. Neural Computation, 1999; 11:229-242.
- [13] Mitrovic D. Reliable Method for Events Recognition. IEEE Trans. on Intelligent Transp. Syst., June 2005; 6(2):198-205.
- [14] Torkkola K, Venkatesan S, Liu H. Sensor Sequence Modeling for Driving. FLAIRS Conference, 2005, Clearwater Beach, Florida, USA, pp. 721-727.
- [15] Zou X. Modeling Intersection Driving Behaviors: A Hidden Markov Model Approach-I. Jrnl. of Transportation Res. Board, 2006: 16-23.
- [16] Boyraz P, Acar M, Kerr D. Signal Modelling and Hidden Markov Models for Driving Manoeuvre Recognition and Driver Fault Diagnosis in an urban road scenario. Proc. Of IEEE IVS'07, 13-15 June 2007, Istanbul, Turkey: 987-992.
- [17] Sathyanarayana A, Boyraz P, Hansen JHL. Driver Behavior Analysis and Route Recognition by Hidden Markov Models. International Conference on Vehicular Electronics and Safety Columbus, September 22-24, 2008, OH, USA.
- [18] Regani SD, Xu Q, Wang B, Wu M, Ray Liu KJ. Driver Authentication for Smart Car Using Wireless Sensing. IEEE Internet Things J., Mar. 2020; 7(3): 2235–2246.
- [19] Fung NC, Wallace B, Chan ADC, Goubran R, Porter MM, Marshall S, Knoefel F. Driver identification using vehicle acceleration and deceleration events from naturalistic driving of older drivers. In 2017 IEEE International Symposium on Medical Measurements and Applications- Proceedings, 2017: 33–38.