IJCTA, 9(2) 2016, pp. 1111-1117

Analysis of Driving Patterns in Car Using Sequential Clustering

Kakde Swati Sudhakar* **and C. Malathy***

ABSTRACT

Recognizing car following behavior of the driver is essential when the mode of transportation is by car as driving is a daily routine for people every time drivers hit the road they face many threats and it is essential to provide a safety and efficient driving.It is proposed to concentration on traffic safety and efficiency which is very important for the driver to drive the car in safe and comfortable manner. Driver behavior is vital to model vehicle dynamics in a traffic simulation environment. Driver behavior depends on many circumstances that are not feasible to collect in driver behavior studies.A new sequential clustering algorithm is developed to find out the driving behaviors and incorporate the comfort interface of the driver which will help to examine driving pattern also. The clustering is based on different data dictionaries such as speed, trip identifier, time, heading, gas pedal position, light intensity; break on off, turn signal state, acceleration etc. that will help to improve the efficiency of car-following behavior as well as driving pattern. Enhance corporate social responsibilities and duty of cares.

Index Terms: Clustering methods, data mining, ITS (Intelligent Transportation System)

I. INTRODUCTION

A driver's behavior is the way a driver reacts to his/her current systems (e.g. speed, distance, accelerate or steer). His/her behavior can, therefore, be formally defined as the function that maps traffic states to a driver's actions. Typically, driver behavior studies collect data that encompass both states and actions and attempts to develop a more accurate mapping between them. These efforts do not usually include many influencing factors of human behavior, such as emotion, personality, hunger, age, gender [18].

The driver's internal factors are Essential in impacting traffic safety in addition to vehicle road environment. The driver's internal factors include the driver psychological characteristics. The differences of driver's age, gender, driving experience and individuality result in different psychological characteristics and the different psychological characteristics are reflected as driving tendencies. The system of driving inclination recognition plays an important role in improving the applicability and accuracy of vehicle active safety systems (collision avoidance as warning system) drivers choose their own desired level of task difficulty [16]. The proposed system is structured such that a known state can be linked to multiple actions thus accounting for the effects of developing a new driving state. A sequential clustering algorithm is developed and used for the clustering of driving behaviors. It is proposed to use sequential clustering algorithm which will segment and cluster a car following behaviors based on different variables such as speed, trip identifier, time, heading, gas pedal position, light intensity; break on off, turn signal state etc. Driver behavior is what the driver chooses to do with these attributes. The objective of this system is to find a perspective set of driving state. Driving is a complicated task. It requires people to see and hear clearly, pay close attention to other cars, traffic signs, signals and react quickly to events. Better driving style and encourage self-management, active driver feedback promotes safety by presenting live feedback and drivers behavior.

* MTech Computer Science & Engineering, SRM University, Kattankulathur, 603203, Tamilnadu, India.

II. CAR-FOLLOWING BEHAVIOR

In recent years, modeling and recognizing driver behavior have become crucial to understanding intelligent transport system, human vehicle systems, and intelligent vehicle system (ITS) [5].Intelligent transportation systems (ITS) are advanced applications which, without embodying intelligence as such, aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated, and 'smarter' use of transport networks. Intelligent transport systems vary in technologies applied, from basic management systems such as car navigation, traffic signal control systems, container management systems, variable message signs, automatic number plate recognition or speed cameras to monitor applications, such as security CCTV systems; and to more advanced applications that integrate live data and feedback from a number of other sources, such as parking guidance and information systems; weather information, bridge de-icing (US deicing) systems.

In the field of transportation and engineering many research and field experiments to study car-following behavior have been conducted on test tracks and roadways, and then modeled to represent driver's behavior [17]. A car the following model was first proposed by Reachel (1950) and Pipes (1953) and was greatly extended by Herman (1959~1967) and has been continuously refined up to present with various approaches to describing relationships between the leader and the following vehicles. Most of the researches into clustering have occurred outside of the field transportation engineering. Largely, there have been two kinds of data collection methods for car-following experiments, using either video recording (or aerial film) to capture many anonymous vehicles, or wire-linked test vehicles on a test track or a roadway [19]. In the first case, it should note that this type of data collection is very difficult and tedious work. One advantage is that real drivers are being observed; who are unaware that the experiment is taking place, thus, their behavior is as "natural" as can be expected.

(A) Data Collection Method

Car-Following studies typically collect vehicle trajectory data through various factors, including naturalistic, simulator and video data collection methods. It also considers the different lane, weather factor, road construction, types of vehicle and traffic flow to analyze it. Data collection mainly observed the drivers physical actions as well as steering movement. The following table gives the description about the different data type and their advantages and flaws including driver's reaction into a different environment. It captures and analyze. The 100 car naturalistic driving study was an instrumented vehicle study conducted in the Northern Virginia/Washington, D.C. area over a two year period. The primary purpose of the study was to collect large scale naturalistic driving data to this end the instrumentation was designed to be discreet, study participants were given no special instructions, and experimenters were not present. Approximately 100 vehicles were instrumented with a suite of sensors including forward and rearward radar, lateral and longitudinal accelerometers, gyro, GPS, access to the vehicle CAN, and five channels of compressed digital video. Collection rates for the various sensors ranged from 1Hz to 10Hz. This collection effort resulted in approximately 2,000,000 vehicles miles and 43,000 hours of driving data

Table I

Data Type	*Description*	*Strengths*	*Weaknesses*
Naturalistic	An instrumental vehicle is driven in normal driving routines	-The driver is in the natural environment. -Multiple trajectories are observed for each driver.	Drivers know they are being observed.

(B) Data Clustering

Cluster analysis is a method of unsupervised learning and a statistical methodology used to categories individual objects into groups with similar meanings (homogeneous) there are several clustering techniques are available within each group, but for the application car-following behavior, the time-series clustering is more suitable. Time series clustering is to partition time series data into groups based on similarity or distance, so that time series in the same cluster are similar. For time series clustering with R, the first step is to work out an appropriate distance/similarity metric, and then, at the second step, use existing clustering techniques, such as k-means, hierarchical clustering, and density-based clustering or subspace clustering, to find clustering structures. Clustering the segments shows how certain behaviors repeat throughout the data within and between drivers.

III. IMPLEMENTATION

Sequential clustering technique is used to capture the full range of state–action clusters for car drivers. First, observed the naturalistic database, followed by extract car-following periods from the huge amount of data. Next, the data is divided into different states to finding a cluster. A car-following period is defined as a period during which one vehicle is reacting to a leading vehicle in the same direction of the car. The behavior is the link between the data set and the action variables that the driver expresses (acceleration, deceleration, lane change, turning). The segmentation process divides each car-following period into multiple segments, and each segment will be defined as a new data set. The sequential clustering process group's similar segments into a single cluster for the new dataset will generate. Thus, a cluster gives the grouping of a certain behavior. Fig 1 shows coordinate system. To compare the different patterns associated with multiple drivers and observing similarities and differences between different drivers.

(A) Naturalistic Driving Data

Naturalistic driving data refer to the data collected from drivers in their natural environment. This is taken by different vehicles with specialized sensors and "vehicle network" recording equipment's, then allowing the drivers to drive the vehicles as they feel comfortable. The equipment records a large number of variables (e.g., speed, acceleration, and steering wheel positions, lane markings), and of particular interest to car following, a radar system positioned at the front of the vehicle records the differences in the position of the car and the speed between the target and leading vehicles. The equipment also includes cameras that record what the driver sees from the front as well as from the two side mirrors, the driver's face, and what the driver is doing inside the vehicle. The naturalistic data used in this system were collected by VTTI. These data were collected from 100 cars that were used by multiple drivers. The 100 Car Naturalistic Driving Study was an instrumented vehicle study conducted in the Northern Virginia / Washington, D.C. area over a two-year period. These data describe crashes or near crashes from the 100 Car Naturalistic Driving Study (Dingus *et al.*, 2006). The data include 68 crashes and 760 near crashes. Each file contains the time series data spanning 30s before an event and 10s after an event. The data in the files is stored as comma delimited text, with each column representing a variable, and each row representing a time sample.

(B) Data Dictionaries

There are different variables used for deriving driver pattern. The variables record will be taken for each individual driver and analyze it. The Table shows the different variables used this variable will help to take a record of driver reacting with those variables.

- List of Dictionary Fields – A description of the components or fields described in the dictionary for each variable entry.

- List of Variables – A list of the entries (variables) in the dictionary which can be used as a table of contents to locate specific variables in the document.
- Conversions, Coordinate System, and Formulas – A catalog of unit conversions, sign conventions and formulas which may be of value to researchers working with these data.
- Data Dictionary Entries – The dictionary entries themselves, one for each variable included in the data set.

The vehicle speed, the lane offset, yaw angle, the range rate represent state variables. The lateral acceleration and the yaw rate represent the driver's action of steering. The longitudinal acceleration represents the driver's actions of accelerating or breaking.

Table II

List of Variables in proposed system

The following variables are included in the text files.

Var	Name
1	Trip Identifier
2	Sync
3	Time
4	Gas pedal position
5	Speed, vehicle composite
6	Speed, GPS horizontal
7	Yaw rate
8	Heading, GPS
9	Lateral acceleration
10	Longitudinal acceleration
11	Lane Markings, Continuity, Left Side Left Line
12	Lane Markings, Continuity, Left Side Right Line
13	Lane Markings, Continuity, Right Side, Left Line
14	Lane Markings, Continuity, Right Side, Right Line
15	Lane Markings, distance left
16	Lane Markings, distance right
17	Lane Markings, type left
18	Lane Markings, type right
19	Lane markings, probability left
20	Lane markings, probability right
21	Radar, forward, ID
22	Radar, rearward, ID
23	Radar, forward, range
24	Radar, rearward, range
25	Radar, forward, range rate
26	Radar, rearward, range rate
27	Radar, forward azimuth
28	Radar, rearward azimuth
29	Light intensity
30	Brake on off
31	Turn signal state

(C) Sequential Clustering

The second part of this methodology clusters the segments, which were found by the segmentation algorithm, in order to find similar segments or behaviors in the car-following data. Clustering the segments shows how certain behaviors repeat throughout the data within and between drivers. Sequential clustering algorithm performs easy computations and processes pattern samples sequentially without special storage requirements. Its behavior is biased by the first patterns passed to the algorithms.

(D) Algorithm

Let d(x, C) denote the distance between data vector x and a cluster C. Furthermore, q is the allowed threshold of dissimilarity. [9]

Let m=1

$C_{m =}X_1$

For i=2 to N,

Find C_k:$d(x_{i,}C_k)$=min $_{i<=j<=m}d(X_{i,}C_j)$

If (d (x_i,C_k)>–”) AND (m<q),

m=m+1

$C_m=x_i$

Else,

$C_k = C_k U x_i$

Where necessary, update representatives

end

end

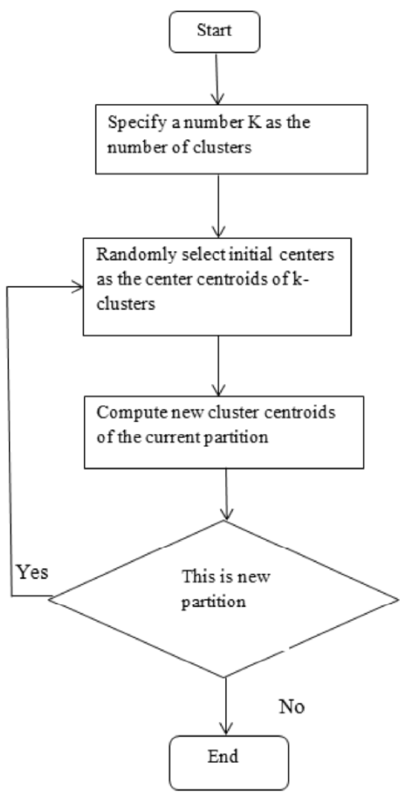

Figure 3: Flow Chart

The algorithm begins choosing the first cluster center among all the pattern samples arbitrarily. Then, it processes the remaining the patterns sequentially. It computes the distance from the actual pattern to its nearest cluster center. If it is smaller than or equal to x, the pattern is assigned to its nearer cluster. If not, a new cluster is formed with the actual pattern. Every m patterns, clusters are merged using a distance criterion (two clusters are combined into one if the distance between their centroids is below C). If there are still more than k clusters, clusters are lumped together using a size criterion (clusters with less than patterns are merged with their nearest clusters). If we still have too many clusters, the nearest pairs of clusters are merged until there are exactly k clusters left.

TABLE II

Cluster	c	m	l	τ (s)	RMSE (m/s)
1	60.07	6.40	7.95	2.95	0.06
2	-42.41	4.00	5.73	3.00	0.69
3	-70.01	-2.39	0.43	3.00	0.78
4	-41.66	-5.25	-2.15	0.20	0.16
5	-4.35	2.34	3.11	3.00	1.13
6	-3.54	0.00	1.00	0.11	1.38
7	-58.81	-6.08	-1.76	2.87	0.74
8	-8.26	-0.24	1.10	1.68	0.61
9	-10.97	0.00	1.00	0.01	1.21
10	-90.72	-3.05	1.00	2.94	1.70
12	-27.02	10.00	9.41	2.91	0.20
13	-0.07	0.52	-0.14	0.00	0.71
14	-22.28	-0.67	0.83	0.06	0.94
15	-12.96	-1.15	0.34	0.62	0.74
16	-27.86	-0.65	1.08	0.01	0.79
17	-18.54	9.18	-3.97	1.93	0.04
19	-31.99	-0.43	1.24	0.53	1.07
20	-72.83	2.86	4.16	3.00	0.63
21	35.27	8.03	10.00	3.00	0.26
23	-21.70	-4.34	-1.85	2.75	0.13
24	38.90	-2.04	1.53	1.31	0.43
25	-1.77	0.51	1.10	1.40	0.80
26	-49.54	-1.70	1.00	2.17	0.95
27	23.76	2.04	7.26	2.58	0.01
28	-94.66	-0.55	1.37	0.08	0.86
29	-72.12	-1.25	0.94	1.28	0.73
30	-16.03	0.06	2.03	1.64	0.79

Sample output

IV. CONCLUSION

It is proposed to find a prescriptive set of state-action clusters that can be used to characterized possible driving patterns of car driver as well as it will help to improve efficiency and safety of the driver. It will help to trained drivers.

REFERENCES

[1] A. Kesting and M. Treiber, “Calibrating car-following models using trajectory data: Methodological study,” Transp. Res. Rec., no. 2088, pp. 148– 156, 2008.

[2] Alejandra Medina Flintsch, Bryan Higgs, Linsen Chong, Montasir M. Abbas, **“**A rule-based neural network approach to model driver naturalistic behavior in traffic” Transportation Research C. Sep-2011

[3] Andrew Liu, NobuyukiKuge, Tomohiro and Osamu Shimoyama **“**A Driver Behavior Recognition Method Based on a Driver Model Framework” Society of Automotive Engineers, 2013

[4] Alex P.Pentland “Driver Behavior Recognition and Prediction in a Smart Car” Media Laboratory Massachustts Institute of Technology (MIT) 20 Ames St E15-385 Cambridge MA02139

[5] B. Higgs, M. Abbas, and A. Medina, “Analysis of the Wiedemann car following model over different speeds using naturalistic data,” presented at the Road Safety Simulation, Indianapolis, Indiana, USA, 2011.

[6] Bryan Higgs and Montasir Abbas, Member, IEEE “Segmentation and clustering of Car-Following Behavior: Recognition of Driving Patterns” Ieee Transactions on Intelligent Transportation Systems, VOL. 16, NO. 1, FEBRUARY 2015

[7] BiagioCiuffo, Carlos Lima Azevedo, JoãoLourenço Cardoso, Moshe E. Ben-Akiva, **“**Dealing with uncertainty in detailed calibration of traffic simulation models for safety assessment” Transportation Research C 2015

[8] Bryan Higgs and Montasir Abbas Ieee transactions on intelligent transportation systems, “Segmentation and Clustering of Car-Following Behavior: Recognition of Driving Patterns”, Member, IEEE.

[9] C. C. Aggarwal and C. K. Reddy, Data Clustering: Algorithms and Applications. Boca Raton, FL, USA: CRC Press, 2013.

[10] ChiyomiMiyajima, FumitadaItakura, Koji Ozawa, Kazuya Takeda, Toshihiro Wakita, Katsunobu Itou, and Yoshihiro Nishiwaki**”** Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification” IEEE 2007.

[11] Christine Buisson**,** LudovicLeclercq, Nicolas Chiabaut. **“**From heterogeneous drivers to macroscopic patterns in congestion” International Journal of Transportation Engineering 2014.

[12] David J. Lovell, David J. Lovell, Gang-Len Chang, Kent L. Norman, Taehyung Kim, William S. Levine, Yongjin Park, **“**Analysis of variability in car-following Behavior over long-term driving maneuvers” Transportation & Environmental Research 2005

[13] David J. Lovell, Martin Hall, Taehyung Kim, Yongjin Park **“**Limitations of Previous Models on Car-Following Behavior and Research Needs “Submitted on the 82th Transportation Research Board Annual Meeting Washington.January -2013.

[14] Ingmar AndréassonandXiaoliangMa**”**Driver reaction time estimation from a real car following data and application in GM-type model evaluation” Center for Traffic Simualtion Research (CTR) Royal Institute of Technology (KTH), Stockholm 10044, Sweden. JingleiZhang, Xiaoyuan Wang “Extraction and Recognition Methods of VehicleDriving Tendency Feature Based on Driver vehicle-environment.

[15] Car following “International Journal of Computational Intelligence Systems, Dec2011

[16] Jiri Macur, Michal Radimsky, TomasApeltauer **“**Validation of microscopic traffic models”http://www.researchgate.net/publication/257465232,April 2013.

[17] M. M. Abbas, B. Higgs, and A. Medina Flintsch, “Naturalistic car driver behavior: Hybrid Wiedemann and Gazis-Herman-Rothery model cali-bration,” presented at the Transportation Research Board 91st Annual Meeting, Washington, DC, USA, 2012.

[18]] N. A. Shenge, S. E. Oladipoand, S. K. Balogun, “Psychosocial factors influencing aggressive driving among commercial and private automobile drivers in Lagos metropolis,” Social Sci. J., vol. 49, no. 1, pp. 83–89, Mar. 2012

[19] T. A. Dingus *et al.* “The 100-car naturalistic driving study; phase II— Results of the 100-car field experiment” Virginia Tech Transp. Inst., Blacksburg, VA, USA, VA DOT HS 810 593, 2006.