



A Developed Prediction Approach for Vehicles Traffic Modeling Based on Regression Models

Roa Sabah Naser¹ , Saad Talib Hasson²

Abstract

Modeling is a safe and efficient approach to solving real vehicle traffic problems. It cooperates with/without simulation to provide feasible methods of analysis, observation, and verification. Modeling can give significant insight into complex systems. Modeling Intelligent Transportation Systems (ITSs) represents a crucial challenge in planning and controlling vehicle traffic congestion. Predicting the vehicle's flow states on roads is the most important challenge in transportation systems. Numerous car-following models have been proposed and developed to illustrate the behavior of moving vehicles. These models are based on real driving assumptions, taking into account velocity and acceleration parameters for each vehicle. The application of car-following models represents an important research direction in enhancing ITSs. In this paper, a car-following model is implemented using a specific vehicle traffic dataset (highD) to predict the vehicle's next velocities after a succeeding period of time based on regression fundamentals. The vehicle's Velocities are based on the driver's behavior. The driver can accelerate, decelerate, or maintain the same speed during any time period of the vehicle journey. Certain threshold values are created by analyzing the recorded real dataset to be used in predicting the next acceleration value for each vehicle at each time period. A regression curve is proposed for each vehicle. From the proposed curves, equations are created to represent the vehicle's velocity. These created equations can be used to predict the vehicle velocity at any given time mathematically.

Keywords: Traffic flow, mobility, car-following models, safety distance model, threshold, regression, simulation, VANET, vehicular communications, Driver behavior

نهج تنبؤ متطور لنمذجة حركة مرور المركبات على أساس نماذج الانحدار

رؤى صباح ناصر¹ ، سعد طالب حسون²

المستخلص

النمذجة هي نهج آمن وفعال لحل مشاكل حركة المركبات الحقيقية. تتعاون مع / بدون محاكاة لتوفير طرق مجدبة للتحليل والمراقبة والتحقق. يمكن أن تعطي النمذجة نظرة ثاقبة للأنظمة المعقدة. تمثل نمذجة أنظمة النقل الذكية (ITS) تحديًا حاسمًا في التخطيط والتحكم في ازدحام حركة مرور المركبات. يعد توقع حالات تدفق المركبات على الطرق من أهم التحديات في أنظمة النقل. تم اقتراح وتطوير العديد من نماذج تتبع السيارات لتوضيح سلوك المركبات المتحركة. تعتمد هذه النماذج على افتراضات القيادة الحقيقية ، مع مراعاة معلمات السرعة والتسارع لكل مركبة. يمثل تطبيق نماذج تتبع السيارة اتجاه بحثي مهم في تعزيز أنظمة النقل الذكية. في هذا البحث ، يتم تنفيذ نموذج تتبع السيارة باستخدام مجموعة بيانات محددة لحركة مرور المركبات (highD) للتنبؤ بالسرعات التالية للمركبة بعد فترة متتالية من المرات بناءً على أساسيات الانحدار. تعتمد سرعات السيارة على سلوك السائق. يمكن للسائق تسريع أو إبطاء أو الحفاظ على نفس السرعة خلال أي فترة زمنية من رحلة السيارة. يتم إنشاء قيم عتبة معينة من خلال تحليل مجموعة البيانات الحقيقية المسجلة لاستخدامها في التنبؤ بقيمة التسارع التالية لكل مركبة في كل فترة زمنية. تم اقتراح منحني انحدار لكل مركبة. من المنحنيات المقترحة ، يتم إنشاء المعادلات لتمثيل سرعة السيارة. يمكن استخدام هذه المعادلات التي تم إنشاؤها للتنبؤ بسرعة السيارة في أي وقت رياضيًا.

الكلمات المفتاحية: تدفق حركة المرور، التنقل، نماذج تتبع السيارات، نموذج مسافة الأمان، العتبة ،

Affiliation of Authors

^{1,2} College of Information Technology, University of Babylon, Iraq, Hilla, 51001

¹roaa.sabah@itnet.uobabylon.edu.iq

²saad_aljebori@itnet.uobabylon.edu.iq

¹ Corresponding Author

Paper Info.

Published: Dec. 2024

انتساب الباحثين

^{1,2} كلية تكنولوجيا المعلومات ، جامعة بابل، العراق، الحلة ، 51001

¹roaa.sabah@itnet.uobabylon.edu.iq

²saad_aljebori@itnet.uobabylon.edu.iq

¹ المؤلف المراسل

معلومات البحث

تاريخ النشر : كانون الاول 2024

Introduction

Vehicular ad-hoc Networks (VANET) are distinguished by their constantly evolving topology, which results from the high speeds of vehicles. The behavior of drivers and surrounding vehicles influences the form of road mobility. The movement patterns of vehicles and their corresponding traffic flow have a significant impact on wireless vehicular communications. [1].

The mobility of the vehicles was modeled to suit the VANET environment simulation based on mathematics, science, and engineering concepts. Researchers defined realistic road conditions for VANET by incorporating traffic constraints (speed, road structure, and vehicle density) with vehicular characteristics and other road constraints. Most of the related works considered vehicle mobility as a significant factor in studying, analyzing, and evaluating the performance of the VANET [2] [3]. The vehicle's arrival rates are considered and can be modeled based on certain probability distribution functions. Most of the previous studies considered the effect of the adjacent vehicles on the traffic pattern of the moving vehicle [4]. The mobility model defined the vehicle's motion during a VANET simulation in a given area. Different mobility models were developed to represent the real behavior of the vehicles on the road [5].

The car-following model was developed to characterize the driving behavior of the following vehicle concerning the leading vehicle in a single lane with overtaking restrictions. It was widely employed in the fields of microscopic traffic simulation, smart driving, traffic safety, and so on [6].

Most of the car-following studies have focused on defining the velocities, acceleration/deceleration values, and their ranges by utilizing different probability distributions [7].

Transportation Engineering focused on estimating the congestion level of roads and their economic and social costs. Different methods were developed to estimate the congestion based on the delay and the required travel time. Changing road conditions are significantly affecting the vehicle's density [8].

Researchers have presented different models for road traffic to analyze vehicle behavior and road conditions. The macroscopic and microscopic models have been considered in different related studies to achieve realistic conditions for moving vehicles on roads. Traffic engineers were focused on density and speed [9].

Related Works

Due to the increase in moving vehicles on roads, researchers suggested different models for road traffic to analyze vehicle behavior and road conditions. The two known models; macroscopic and microscopic models have been considered in different studies to achieve realistic conditions to deal with moving vehicles on roads. The focus of traffic engineers is on density, speed, and drivers' behavior [10].

Researchers in [11] developed a car following model. They integrated the heterogeneity of the driver's responsiveness using a proposed function of the optimal velocity with the prior car-following model. They used a linear analysis method to determine the revised model's stability criterion. Their suggested model was numerically simulated, and its result indicated a significant influence on

enhancing traffic stability and lowering traffic congestion.

A proposed traffic flow model based on the full velocity difference model was developed by [12]. The authors tried to examine the backward-looking effect (forward-backward velocity difference) model. They adopted a modified backward optimal velocity and generalized backward maximum speed.

Authors in [13] proposed a method to validate the Safety distance model. Their calibration approach begins with validating the steady-state car-following model with data from macroscopic loop detectors. They calibrated the vehicle acceleration component using car specification data acquired from automobile manufacturers. The light-duty and heavy-duty vehicle field data were utilized to validate their calibration processes.

Car-Following Models

Car-following models have been widely used in traffic flow simulations to predict the behavior of individual vehicles on the road. One of the earliest car-following models was the safety distance model, which was proposed by Gazis et al. in 1961 [6].

The safety distance model assumes that vehicles maintain a safe distance from the vehicle in front of them, which is proportional to their speed. The model defines the safe distance as the sum of the perception-reaction time and the stopping distance of the vehicle. The perception-reaction time is the time taken by the driver to perceive and react to a change in the front vehicle's behavior, while the stopping distance is the distance required to stop the vehicle at a given speed [11]. The micro-simulation tool defines the maximum, minimum, and desirable acceleration/deceleration values, maximum, desired speed values, and other vehicle parameters for any car-following model using a variety of statistical distributions and functions [12].

The safety distance model represents one of the earliest car-following models proposed and widely used in traffic flow simulations. While it has been extended and modified over the years, it has some limitations, and newer car-following models have been developed to overcome these limitations. The acceleration sub-model is corresponding to the empirical formulations shown by equation (1) [12].

$$v_n^{acc}(t + \Delta t) = v_n(t) + 2.5 \cdot a_n \cdot \Delta t \left[1 - \frac{v_n(t)}{\text{desired speed}} \right] \sqrt{0.025 + \frac{v_n(t)}{\text{desired speed}}} \tag{1}$$

Where the acceleration of the vehicle is $v_n^{acc}(t + \Delta t)$, Δt is the reaction time, $v_n(t)$ is the speeds of vehicles n at time t , *desired speed* is the desired speed that the following vehicle wants to reach, and a_n is the maximum acceleration for a vehicle n .

The traditional safety distance car-following model is shown in Equation (1). In this model the value of the velocity and acceleration for the vehicle (n) at the time (t) is used to predict its next velocity at the time (t + Δt). Usually, the value of the velocity will be adapted to the acceleration value. An increase in acceleration will result in an increase in velocity, while a decrease in acceleration (deceleration) will result in a decrease in velocity.

Improved Safety Distance Model

The fixed value of acceleration will result in a fixed velocity.

A modification of equation (1) is proposed in this research paper. A new variable (z) is proposed to control the sign of the acceleration factor value in equation (1). The value of z will be (1, -1, or 0). When the relative acceleration (difference between the current acceleration value and the previous acceleration value) is positive, then the value of $z = 1$. This case will result in an increase in the vehicle’s velocity due to its acceleration. When the

$$v_n^{acc}(t + \Delta t) = v_n(t) + 2.5 \cdot z \cdot a_n \cdot \Delta t \left[1 - \frac{v_n(t)}{\text{desired speed}} \right] \sqrt{0.025 + \frac{v_n(t)}{\text{desired speed}}} \tag{2}$$

Where the acceleration of the vehicle is $v_n^{acc}(t + \Delta t)$, Δt is the reaction time, $v_n(t)$ is the speeds of vehicles n at time t , *desired speed* is the desired speed that the following vehicle wants to reach, and a_n is the maximum acceleration vehicle n .

This modified equation can be used in developing the simulation experiments to predict the next future velocity for each vehicle at each time.

Data Analysis

In order to implement and evaluate the modified model as a car-following model, a highD dataset is utilized. It was downloaded from “HighD Dataset (highd-dataset.com)“. The highD dataset is a brand-new collection of realistic vehicle trajectories captured on German roads. It recorded 110500 vehicles, with a totally driven distance of 44500 kilometers, and 147 driven hours at six distinct sites traffics. Although the dataset was intended for the safety validation of highly automated cars, it may also be used for other purposes such as traffic pattern analysis and driver model modeling.

The dataset features are considered and arranged carefully. The first step is to isolate the velocities

relative acceleration (difference between the current acceleration value and the previous acceleration value) is negative, then the value of $z = -1$. This case will result in a decrease in the vehicle’s velocity due to its deceleration. When the relative acceleration (difference between the current acceleration value and the previous acceleration value) is zero, then the value of $z = 0$. This case will result in keeping the vehicle at its previous velocity. The proposed improved equation can be rewritten in equation (2).

and the acceleration values for the moving vehicles in each direction. One direction road is called the east, and the opposing direction is called the west direction road in this study analysis.

1- Velocity and Acceleration Statistics

Each vehicle in the file (car or truck type) has different velocities and acceleration values in different recording frames. The positive velocity values are assigned to the moving vehicles on the east-direction road and the vehicles with negative velocities are assigned to the moving vehicles on the west-direction road.

The vehicle’s velocities are recorded directly from the dataset records (they were observed practically). Each vehicle has one velocity and acceleration value in each frame. The dataset records the acceleration and the velocity for each vehicle at a different number of frames. The mean velocity, mean acceleration, and standard deviations for each vehicle are computed. Equations (3 and 4) are utilized in computing these means.

$$avg_{v_k} = \frac{\sum_{i=1}^n v_i}{n} \tag{3}$$

$$avg_{acc_k} = \frac{\sum_{i=1}^n acc_i}{n} \tag{4}$$

Where v_i , is the recorded vehicle’s velocity in frame i , n is the number of frames, avg_{v_k} is the average velocity for the vehicle k , avg_{acc_k} is the average acceleration for vehicle k and acc_i is the acceleration of a vehicle in a frame i .

2- Threshold Estimation

For each road (east and west), vehicles’ average acceleration values are sorted, analyzed, and arranged as an accelerated case (positive values), a decelerated case (negative values), and neither an acceleration nor a deceleration case (zero values). These three cases are observed and counted. Table (1) shows the total collected results for both roads.

Table (1): vehicles acceleration and deceleration cases

Acceleration value	East file	West file
Number-of-Positive Values	313	277
Number-of-Negative Values	113	215
Number-of-Zero Values	28	103
total	454	595

Table (1) is used to estimate certain fractions. These fractions are applied as certain probability values. The estimated probability values are proposed as threshold values to be used for

predicting the next time acceleration for each vehicle in implementing a simulation experiment. Table (2) shows these proposed threshold values.

Table (2): Proposed Threshold values.

Probability	East value	West value
Probability of increased acceleration	0.69	0.47
Probability of decreased acceleration	0.25	0.36
Probability of fixed acceleration	0.06	0.17

Table (2) shows that the drivers' increased acceleration in both directions is high compared with the deceleration behavior. Driving at a fixed velocity without changing the acceleration is about 10%.

To proceed with the dataset information, a prediction model is developed and implemented based on the fundamentals of the Markov model.

This developed model is based on creating sequenced acceleration values after each short time period by utilizing equation 2 to create the new velocities. Numerical results are estimated and recorded to show the implementation process based on the modified equation. A simple modeling and simulation process in implemented to estimate the next four velocity values.

A simulation scenario is implemented by generating a random number for each step of the prediction to test the next behavior based on the comparison between the generated random number and the created threshold values. Certain limits are created to generate specific categories for each case based on the generated threshold values. These categories are 0 - 0.69, 0.691- 0.75, and 0.751- 1.0 for the East road direction and 0- 0.47, 0.471- 0.64, and 0.641- 1.0 for the West road direction.

For the east road, if the generated random number falls in the range of (0 - 0.69), the vehicle will tend to accelerate, so the value of the variable z in equation 2 will be 1. If the value of the generated random number falls between 0.691- 0.75, then the acceleration will not change, so the value of the variable z will be 0. If the value of the random number is between (0.751 – 1), then the vehicle will decelerate and the value of z will be -1.

For the west road, if the generated random number falls in the range of (0 - 0.47), the vehicle will tend

to accelerate, so the value of the variable z will be 1. If the value of the random number falls between 0.471- 0.64, then the acceleration will not change, so the value of the variable z will be 0. If the value of the random number is between (0.641– 1), then the vehicle will decelerate and the value of z will be -1.

Four steps prediction values are performed to estimate the vehicle’s velocity in each time period. Equation 2 is implemented to estimate the next time velocity of a vehicle based on its current velocity. Threshold categories and random numbers are utilized in these calculations. Tables (3 and 4) present the simulation experiments to estimate the vehicle’s velocities on the west and east roads after certain time periods. These two tables present the next estimated velocities for the vehicles on the west and east roads. These velocities are estimated after a certain time $\Delta t = 5s$, $2 \Delta t = 10s$, $3 \Delta t = 15s$, and $4 \Delta t = 20s$.

Table (3): west street velocity prediction

AVG(V+)	Step1(v)	Step2(v)	Step3(v)	Step4(v)
31.9861538	31.8371043	31.98615	32.17673884	32.2714952
42.7624223	42.6889942	42.75546	42.7554629	42.762428
43.3986813	43.3309864	43.45947	43.2765993	43.399177
31.3115228	31.1527279	31.33054	31.3115361	31.273475
36.4210762	36.2959959	36.43740	36.4219184	36.437446
38.8742857	38.767969	39.13248	38.8792625	38.879262
37.9635424	37.8458719	37.99361	37.9636040	37.903342
44.4743722	44.4168725	44.50938	44.4743722	44.439559
30.7895195	30.6277159	30.70130	30.7897921	30.701302
24.500645	24.309689	24.51428	24.500649	24.527917

Table (4): east street velocity prediction

AVG(V)	Step1(v)	Step2(v)	Step3(v)	Step4(v)
--------	----------	----------	----------	----------

41.071212	41.183260	41.183260	41.294093	41.184466
36.130828	36.192279	36.192279	36.192279	36.192494
23.328	23.339934	23.339934	23.351866	23.363796
33.800923	33.941294	33.941294	34.080757	34.211218
22.780818	22.699539	22.780818	22.699539	22.536696
24.786340	24.925656	24.647023	25.064600	25.064600
35.568621	35.450125	35.687118	35.330872	35.450883
42.137926	42.265748	42.264833	42.139747	42.139747
35.519103	35.488285	35.549921	35.457416	35.488336
36.538581	36.572809	36.504353	36.606967	36.572878

Tables 3 and 4 show four-step prediction periods for each vehicle. These results reflect the randomness and independent case for each value, which is very close to the real road velocity. Each predicted velocity value may increase or decrease compared with its previous value.

Models Results Evaluation

To perform certain evaluation processes, a goodness-of-fit approach is applied to show the validity of the results in Tables (3 and 4). The Chi-square test as a goodness of fit tool is applied to check if there is any difference between the

observed velocity values (measured by the dataset) and the expected velocity value calculated by the developed prediction approach. The formula for chi-square can be written as [14]:

$$\chi^2 = \sum \frac{(Ov_i - Ev_i)^2}{Ev_i} \tag{5}$$

Where Ov_i , and Ev_i represent the observed velocity and expected velocity values, respectively. Tables (5 and 6) present the Chi-Square values for tables (3 and 4) respectively. These values are compared with the tabular accepted values for the Chi-Square value.

Table (5): chi-square for West Street

AVG(V+)	Step1(v)	Step2(v)	Step3(v)	Step4(v)
0.2586	0.4307	0.5447	0.79	

Table (6): chi-square for West Street

AVG(V-)	Step1(v)	Step2(v)	Step3(v)	Step4(v)
0.18	0.1735	0.4041	0.6348	

From Tables (5 and 6), the calculated values for the Chi-square values are accepted and the results are trusted compared with the standard tabular values.

Regression

Regression is a statistical method used to analyze the relationship between a dependent variable and one or more independent variables. The goal of regression analysis is to identify the strength and

direction of the relationship between the variables, as well as to predict the value of the dependent variable based on the values of the independent variables. [15]. To make use of the developed results with the improved safety distance mobility model, this paper uses regression models as a tool to perform a long-time prediction to predict the vehicle velocity in the next given time based on the available initial velocity at the starting point.

The velocity results in Tables (3 and 4) are used in developing a regression curve for each vehicle. The regression curve for each vehicle is fitted, and its parameters are estimated based on the calculated values. Figures 1, 2, 3, and 4 show samples of these proposed regression curves for certain vehicles.

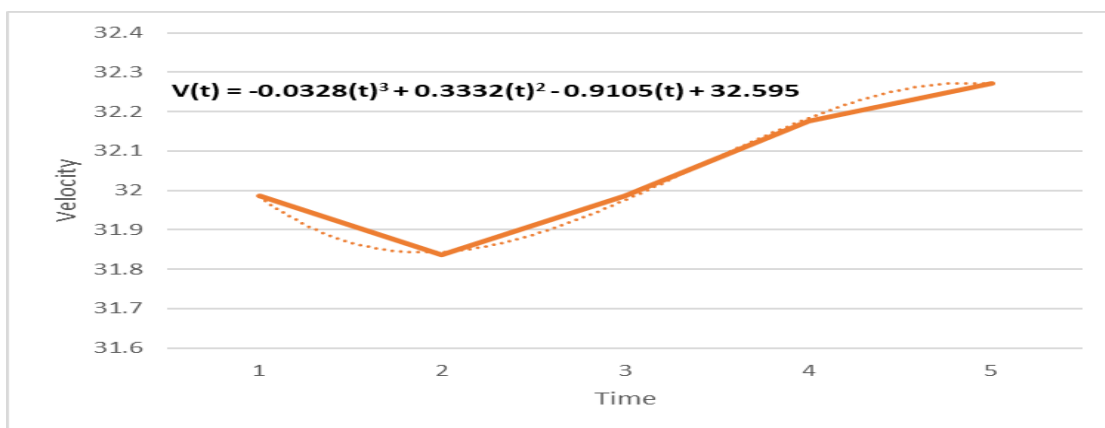


Figure (1): State the Regression curve for vehicle (1) from the west direction street

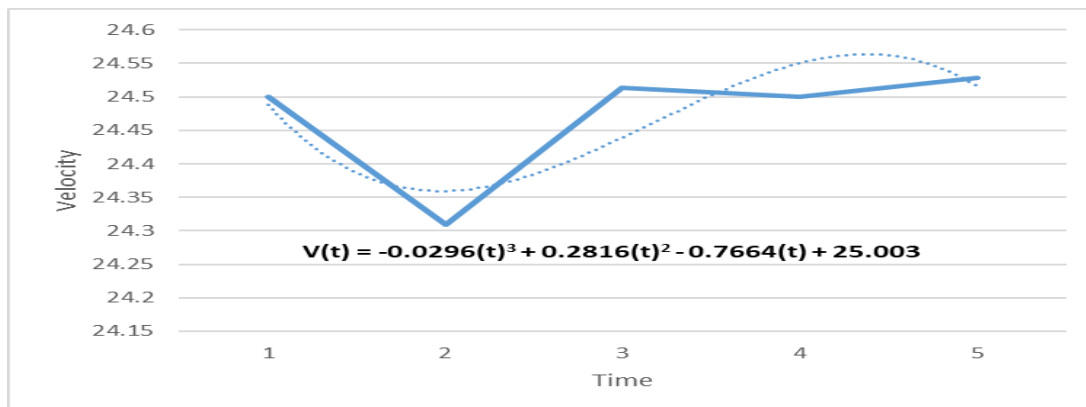


Figure (2): State the Regression curve for vehicle (10) from the west direction street

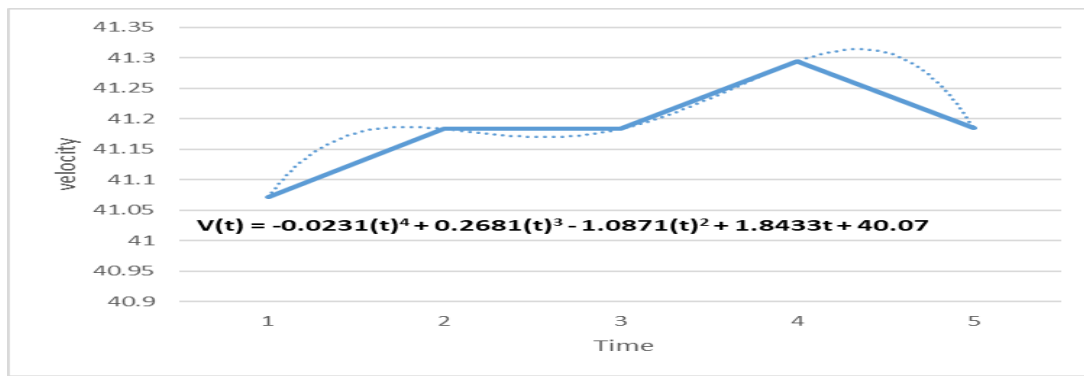


Figure (3): State the Regression curve for vehicle (1) from the east direction street

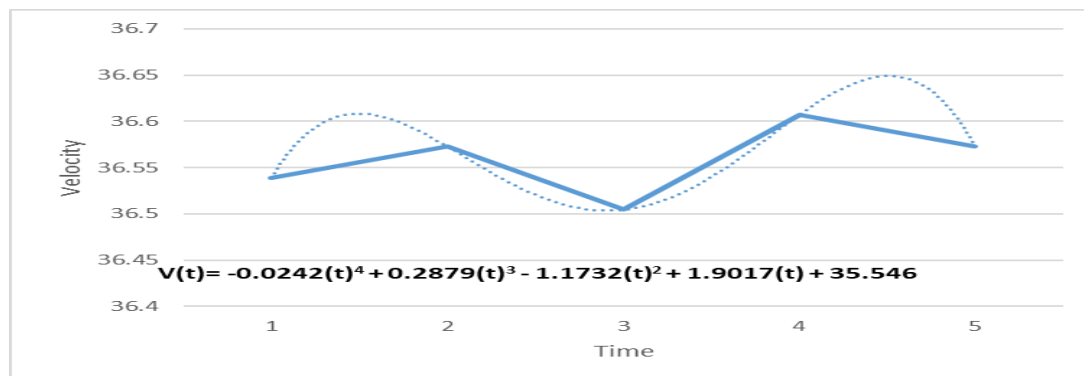


Figure (4): State the Regression curve for vehicle (10) from the east direction street

From the proposed regression models, a developed equation can be created for each vehicle. A vehicle velocity can be estimated from its created equation at each time (t) directly. Such an approach will help in predicting the vehicle’s velocities in a simple mathematical manner without requiring detailed analysis.

Conclusion

Car-following models represent a considerable investigation field to analyze and evaluate the roads traffic. Observing and controlling the traffic flow on roads can be predicted, assisted, and analyzed by applying current or developed car-following models. In this paper, the car-following behavior was examined and analyzed with consideration of certain driver behaviors in a real dataset (HighD). A modification is added to the original model to create a valid car-following

model. A dataset is used in implementing the developed model. Predicting the vehicle’s behavior (accelerate, decelerate, or not) at each time is performed based on the added modification variable. Such a prediction model is proposed and implemented to predict the next time velocity and acceleration for the moving vehicles. Using such models will help in avoiding or controlling traffic congestion and reduce road accidents. A regression model is implemented as another tool in this paper. A regression curve can be fitted for each vehicle’s velocities at different times. From the regression curve, a proposed equation is created to represent the vehicle’s velocities at each time. The created equation is used simply to find the vehicle’s velocity at each time (t).

References

- [1] T. Umer, M. Amjad, N. Shah, and Z. Ding, “Modeling Vehicles Mobility for Connectivity Analysis in VANET,” *Intelligent Transportation Systems*, pp. 221–239, 2016, DOI: 10.1007/978-3-319-28183-4_10.
- [2] L. Du and S. Ukkusuri, “The Relative Mobility of Vehicles Improve the Performance of Information Flow in Vehicle Ad Hoc Networks,” *Networks and Spatial Economics*, vol. 10, no. 2, pp. 209–240, May 2008, DOI: 10.1007/s11067-008-9063-x.
- [3] H. V. and V. P*, “Enhancing Mobility Based on Realistic Models for Vehicular Ad Hoc Networks,” *International Journal of Engineering and Advanced Technology*, vol. 10, no. 4, pp. 220–227, Apr. 2021, DOI: 10.35940/ijeat.d2503.0410421.
- [4] S. Ur Rehman, M. Khan, T. Zia, and L. Zheng, “Vehicular Ad-Hoc Networks (VANETs) -An Overview and Challenges,” *Journal of Wireless Networking and Communications*, vol. 2013, no. 3, pp. 29–38, DOI: 10.5923/j.jwnc.20130303.02.
- [5] J. Härri, F. Filali, and C. Bonnet, “Mobility models for vehicular ad hoc networks: a survey and taxonomy,” undefined, 2009, Accessed: Nov. 12, 2022. [Online]. Available: <https://www.semanticscholar.org/paper/Mobility-models-for-vehicular-ad-hoc-networks%3A-a-H%C3%A4rri-Filali/522577dd187ac2f9683751d5cf8995e0275cff92>
- [6] F. J. Ros, J. A. Martinez, and P. M. Ruiz, “Mobility Models, Topology, and Simulations in VANET,” *Mobile Ad Hoc Networking*, pp. 545–576, Mar. 2013, DOI: 10.1002/9781118511305.ch15.
- [7] R. Fu, Z. Li, Q. Sun, and C. Wang, “Human-like car-following model for autonomous vehicles considering the cut-in behavior of other vehicles in mixed traffic,” *Accident Analysis & Prevention*, vol. 132, p. 105260, Nov. 2019, DOI: 10.1016/j.aap.2019.105260.
- [8] T. Afrin and N. Yodo, “A Survey of Road Traffic Congestion Measure towards a Sustainable and Resilient Transportation System,” *Sustainability*, vol. 12, no. 11, p. 4660, Jun. 2020, doi: 10.3390/su12114660.
- [9] H. K. Lee, H.-W. . Lee, and D. Kim, “Macroscopic traffic models from microscopic car-following models,” *Physical Review E*, vol. 64, no. 5, Oct. 2001, DOI: 10.1103/physreve.64.056126.
- [10] H. K. Gaddam and K. R. Rao, “Modelling vehicular behavior using trajectory data under non-lane based heterogeneous traffic conditions,” *Archives of Transport*, vol. 52, no. 4, pp. 95–108, Dec. 2019, DOI: 10.5604/01.3001.0014.0211.
- [11] “Car-following.” Accessed: Nov. 12, 2022. [Online]. Available: <https://ocw.tudelft.nl/wp-content/uploads/Chapter-8-Car-following.pdf>
- [12] H. U. Ahmed, Y. Huang, and P. Lu, “A Review of Car-Following Models and Modeling Tools for Human and Autonomous-Ready Driving Behaviors in Micro-Simulation,” *Smart Cities*, vol. 4, no. 1, pp. 314–335, Mar. 2021, DOI: 10.3390/smartcities4010019.
- [13] L. Vasconcelos, L. Neto, S. Santos, A. B. Silva, and Á. Seco, “Calibration of the Gipps Car-following Model Using Trajectory Data,” *Transportation Research Procedia*, vol. 3, pp. 952–961, Jan. 2014, DOI: 10.1016/j.trpro.2014.10.075.

- [14] S. Kilic, "Chi-square Test," *Journal of Mood Disorders*, vol. 6, no. 3, p. 180, 2016, DOI: 10.5455/jmood.20160803110534.
- [15] S. Kallam, R. Patan, T. V. Ramana, and A. H. Gandomi, "Linear Weighted Regression and Energy-Aware Greedy Scheduling for Heterogeneous Big Data," *Electronics*, vol. 10, no. 5, p. 554, Feb. 2021, doi: 10.3390/electronics10050554.