



## Article

# Bus Travel Time Prediction Based on the Similarity in Drivers' Driving Styles

Zhenzhong Yin <sup>1,\*</sup> and Bin Zhang <sup>2,\*</sup><sup>1</sup> School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China<sup>2</sup> Software College, Northeastern University, Shenyang 110819, China

\* Correspondence: yinzhenzhong@stumail.neu.edu.cn (Z.Y.); zhangbin@mail.neu.edu.cn (B.Z.)

**Abstract:** Providing accurate and real-time bus travel time information is crucial for both passengers and public transportation managers. However, in the traditional bus travel time prediction model, due to the lack of consideration of the influence of different bus drivers' driving styles on the bus travel time, the prediction result is not ideal. In the traditional bus travel time prediction model, the historical travel data of all drivers in the entire bus line are usually used for training and prediction. Due to great differences in individual driving styles, the eigenvalues of drivers' driving parameters are widely distributed. Therefore, the prediction accuracy of the model trained by this dataset is low. At the same time, the training time of the model is too long due to the large sample size, making it difficult to provide a timely prediction in practical applications. However, if only the historical dataset of a single driver is used for training and prediction, the amount of training data is too small, and it is also difficult to accurately predict travel time. To solve these problems, this paper proposes a method to predict bus travel times based on the similarity of drivers' driving styles. Firstly, the historical travel time data of different drivers are clustered, and then the corresponding types of drivers' historical data are used to predict the travel time, so as to improve the accuracy and speed of the travel time prediction. We evaluated our approach using a real-world bus trajectory dataset collected in Shenyang, China. The experimental results show that the accuracy of the proposed method is 13.4% higher than that of the traditional method.

**Keywords:** bus travel time prediction; driving style; hierarchical clustering; machine learning; support vector machines



**Citation:** Yin, Z.; Zhang, B. Bus Travel Time Prediction Based on the Similarity in Drivers' Driving Styles. *Future Internet* **2023**, *15*, 222. <https://doi.org/10.3390/fi15070222>

Academic Editor: Wei-Chiang Hong

Received: 23 May 2023

Revised: 14 June 2023

Accepted: 19 June 2023

Published: 21 June 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the increasing number of urban vehicles, traffic congestion and air pollution have become major problems in many cities. One of the important ways to solve urban traffic problems is to prioritize the development of public transportation and reduce the number of private cars used by commuters. The key to the development of public transport is providing good transport services for passengers. It is very important to provide accurate bus travel time information to passengers, which can help travelers plan their trips and reduce waiting time [1,2]. However, due to the complex urban transportation environment, accurately predicting bus travel time is very difficult. Multiple factors can affect the travel time of buses, such as changes in passenger flow, traffic conditions, the time period of operation, the driver's driving style, weather factors, etc. [3,4]. These factors lead to great uncertainty in bus travel times [5].

Due to a lack of sufficient data, some of these influencing factors are difficult to quantify, and it is difficult to establish a model that includes all of the factors that affect travel time [6,7]. For example, existing research does not consider the influence of differences in drivers' individual driving styles on bus travel times. Because each bus is driven by different drivers, the level of driving, driving experience, and driving habits of each driver are different, which has a significant impact on the travel time of buses [3,7]. For example,

when passing the same road section, a driver with an aggressive driving style will likely arrive at the next stop earlier than the scheduled time, a driver with a steady driving style will likely arrive at the next stop on time, and a driver with an overly gentle or cautious driving style will very likely arrive at the next stop later than the scheduled time. Therefore, the difference in travel time caused by differences in driving style increases as the distance traveled increases. As a result, a bus that departs later will gradually catch up with a bus that leaves earlier, resulting in a crowding of buses. This phenomenon is called bus bunching [8], as shown in Figure 1, supposing that the driving style of bus driver a is too gentle, the driving style of bus driver b is more steady, and the driving style of bus driver c is more aggressive. Their departure times are 10 min apart, but differences in driving styles lead them to meet at the 12th stop. This phenomenon disrupts the regularity of bus operations. Regularity is the main indicator of service reliability since it is the main determinant of passenger waiting times [9]. Bus bunching is not conducive to passengers being able to the bus on time and affects the enthusiasm of passengers towards taking the bus. With the rapid development of GPS positioning technology, big data technology, and machine learning technology, the differences in bus travel time caused by drivers' individual driving behaviors can be found regularly as long as sufficient data are obtained.

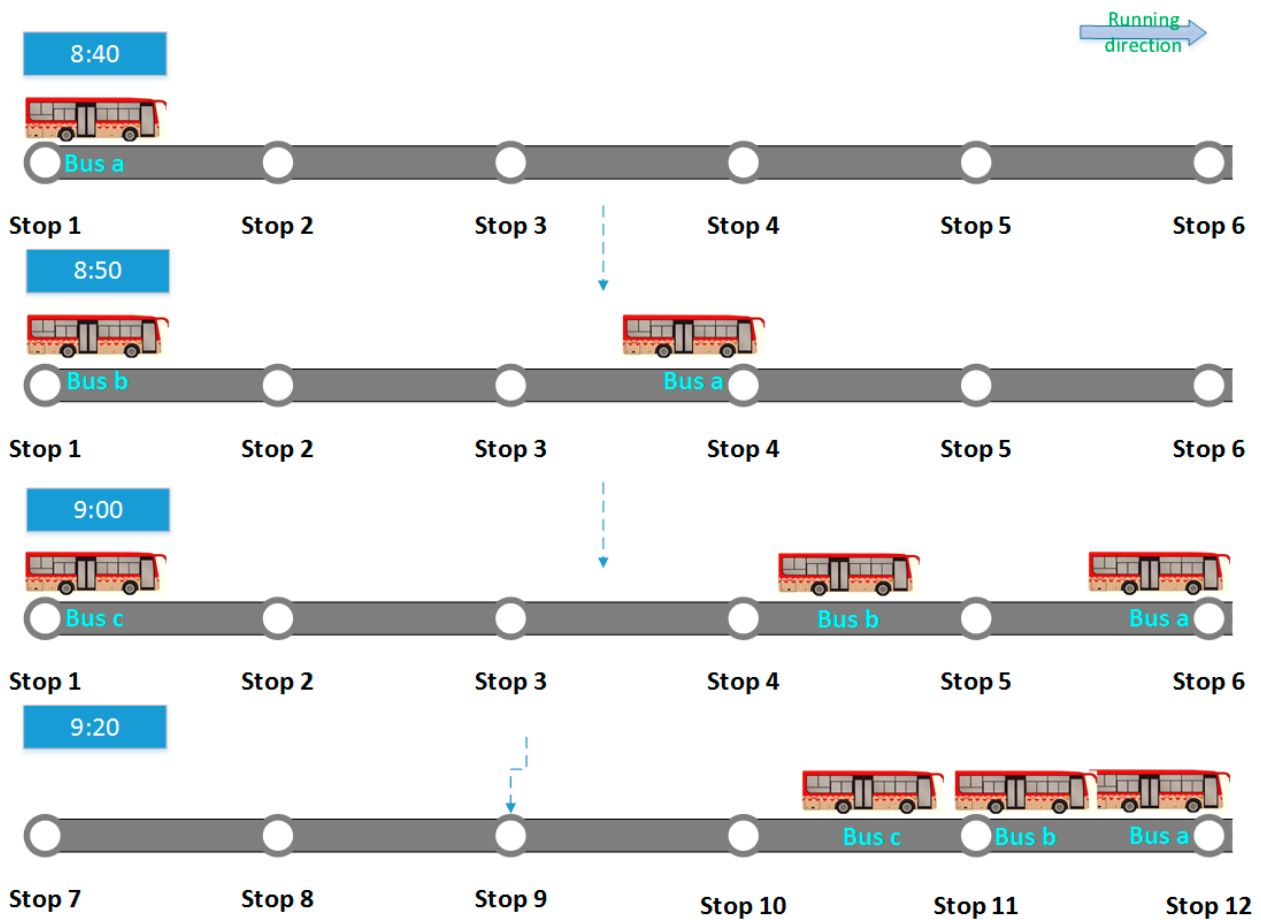


Figure 1. Bus bunching illustration.

This article aims to introduce the factor of the driver's driving style into the bus travel time prediction model through a study of the driving styles of city bus drivers. This is helpful for bus operators to promptly correct the impact of drivers' styles of driving on travel time according to the prediction results, thus reducing the bus bunching phenomenon. To effectively distinguish the driving styles of different drivers, this article first introduces the bus driver driving style judgment model based on hierarchical clustering technology, groups the historical travel time data of different drivers, and then uses the corresponding travel time prediction model to predict bus travel time.

The main contributions of this paper can be summarized as follows.

- We propose a new combination method to study the influence of driving style factors on the prediction of travel time;
- To our knowledge, this is the first study to introduce the factor of the driver's driving style into the bus travel time prediction model;
- We are able to improve the accuracy of travel time forecasting and shorten the forecasting time. Our preliminary results reveal the influence of drivers' driving styles on bus travel times.

This paper is organized as follows. Section 2 presents a brief review of the background literature. Section 3 includes the classification method of the driving styles of bus drivers and the prediction model of bus travel time. In Section 4, the effectiveness of the proposed method is verified by experiments, and the experimental results are analyzed and discussed. Section 5 is the conclusion and a brief note on future research directions.

## 2. Related Works

The prediction of travel time is an essential and problematic component of the Intelligent Public Transport System (IPTS) [10] that has attracted many researchers and IPTS planners [7]. Various models have been proposed to predict bus travel times or arrival times. Some studies have also mentioned that bus travel time is affected by the driving styles or driver behaviors of bus drivers [7,11]. However, due to a lack of data, no in-depth research studies have been performed. This article will review existing research in terms of two aspects: the classification of driver driving styles and bus travel time prediction.

### 2.1. Research on Driver's Driving Style

A driver's driving style can be understood as the way a driver operates the vehicle controls in driving scenarios and external conditions [12]. Driving style is largely dependent on traffic conditions, vehicle performance, the driver's personality, and control devices at traffic intersections [13]. Driving style is manifested mainly in a driver's control over the vehicle's speed and acceleration, as well as their lane change behavior and progress during driving [14]. For analyses of driving behavior, it is important to select appropriate test cycles, test routes, test vehicles, and test drivers [14]. Vehicle positioning can be provided by GPS, which is an indirect measurement method of speed and acceleration at the same time.

The most common aggregate driving behavior parameters used in the articles studied include the mean speed, acceleration, deceleration, driving duration, number of acceleration and deceleration events, relative positive speed, and idle time [14]. Speed and acceleration/deceleration were used by Johnson and Trivedi to identify driver characteristics [15]. Murphey et al. [16] categorized drivers' driving styles according to the speed at which a driver accelerated and decelerated. Moreover, they proposed that driving style is a short-lived behavior; specifically, a driver can be aggressive over a period of time, while their performance might be normal during another period of time. Langari and Won [17] used standard deviation and the average acceleration ratio to classify driving styles. If the ratio was greater than 100%, the driving style was classified as aggressive, if the ratio was between 50% and 100%, the driving style was classified as normal, and if the ratio was less than 50%, the driving style was classified as calm.

The classification algorithms for the driving styles of drivers are usually divided into three categories in these papers. The first category is those that are implemented by a set of

rules, also known as threshold-based algorithms. It is the simplest method for recognizing driving styles [17,18]. This type of algorithm is simple to use, as well as easy to explain and implement, but it limits the number of parameters that can be managed. Therefore, its accuracy is quite limited. The second category is model-based classification algorithms, which describe driving styles through a set of predefined feature equations. However, the main disadvantage of driving style modeling by drivers is that it is difficult to prove the accuracy of the results. For model verification, its results must be compared with those of real drivers, which requires a lot of data collection [15,19]. The third category is recognizing drivers' driving styles through machine learning algorithms, including the hierarchical cluster analysis [20], analysis of principal components [20], Gaussian mixture model [19], k-nearest neighbor [21,22], artificial neural networks [22,23] and other machine learning algorithms.

Current research generally divides the driving style of drivers into two or three categories. In [15,18,21,22], driving style was divided into two categories, i.e., aggressive and non-aggressive. Xu et al. [24] divided driving styles into three categories, including aggressive, mild, and medium driving styles. Murphey et al. [16] divided driving styles into four categories by analyzing the jerk profile of the driver, including calm driving, normal driving, aggressive driving, and no speed. Z. Constantinescu et al. [20] proposed five to seven levels of drivers, covering a range from non-aggressive to aggressive. However, this complicated the algorithm's development and the interpretation of the classes themselves. Augustynowicz [23] classified drivers using a range within  $(-1, 1)$ , from the mild driver to the aggressive driver. Despite a fruitful line of empirical findings from previous scholarship, the impact of bus drivers' driving styles on travel time has not yet been fully investigated.

## 2.2. Research on Prediction of Bus Travel Time

In previous studies, various complex models and algorithms have been developed to predict bus travel time by using AVL/APC/GPS data. In recent years, with the rapid growth of data and the expansion of Intelligent Transportation Systems (ITS), machine learning has become an important tool for solving complex problems, such as the prediction, analytics, and patterns of large amounts of data [25]. The following models are mainly used for applying machine learning to predict the travel time of buses.

Multivariate linear regression models use multivariate statistical techniques to examine the linear correlation between a group of independent variables and a single dependent variable [1]. Patnaik et al. [26] used multiple linear regression models to predict the arrival time of buses at a target bus stop. This kind of model can reflect which independent variables are important for travel time prediction and which independent variables are relatively less important for travel time prediction. This model has a small number of calculations and can get satisfactory results in some relatively simple cases. However, the variables in a transportation system are interrelated, so the applicability of the regression model is generally limited [27].

An artificial neural network (ANN) is produced by simulating the intelligent data processing ability of the human brain. Due to its outstanding advantages in solving complex non-linear problems, ANNs are very popular for travel time prediction [28,29]. An ANN model was used to predict bus arrival times in Jeong and Rilett [1]. Their results show that the ANN model is better than the regression model in terms of prediction accuracy. Yu et al. [30] concluded that although the performance of ANN models is worse than that of SVMs, ANN models are better than k-NN and LR models. Fan and Gurmu [2] used only GPS data to predict bus travel times and concluded that the ANN model is superior to the historical average (HA) model and the Kalman filter model in terms of its overall accuracy and the robustness of its prediction.

The support vector machine (SVM) model is similar to an ANN. An SVM has a strong learning ability and better generalization ability than a neural network. It is easy to balance the degree of fitting and the level of generalization. It shows many unique advantages in solving small-sample, nonlinear, and high-dimensional pattern recognition problems [31]. Wu et al. [31] first applied a support vector regression model to the prediction

of travel time on a highway, proving that the SVR is suitable for traffic data analyses and demonstrated the feasibility of applying SVRs in the prediction of travel time. Vanajakshi and Rilett [32] compared several different travel time prediction methods, including a historical method, time series analysis, ANN, and SVM. Their results show that the SVM is a feasible alternative to a short-term prediction problem when the amount of data is essentially small or noisy. Yu et al. [33] used an SVM to predict bus arrival times, to study the feasibility and applicability of SVMs in the field of predicting bus travel times. Then Yu et al. [30] compared the SVM, ANN, k-NN algorithm, and LR regression models. Their results show that the SVM model is the most accurate model among the four models. Ma et al. [7] compared the ANN, KNN, and SVM bus travel time prediction models through experiments, and their results show that the performance of the SVM is better than that of the other two algorithms.

The Kalman filter (KF) is a linear recursive prediction updating algorithm, which is used to estimate the parameters of the process model. By using dynamic AVL (Automatic Vehicle Location) and APC (Automatic Passenger Counting) data, Shalaby and Farhan [34] tried to use the KF bus travel time model to provide real-time information on the arrival and departure times of buses. Vanajakshi et al. [35] used the KF model to predict bus travel times, and their conclusion showed that the KF model was significantly better than the average method. Fan and Gurmu [2] used a historical average model, KF model, and ANN model to predict the travel times of buses. They concluded that the prediction effect of the KF model was better when there were no huge differences in the travel time between long-distance road sections and two adjacent sections. In two other studies [28,36], a KF model was used as a dynamic adjustment algorithm to correct the baseline travel time predicted by an ANN and SVM.

Throughout our review of previous studies, SVMs have shown better performance in predicting bus travel times compared to other models. Furthermore, they are advantageous in solving small-sample, nonlinear, and high-dimensional pattern recognition problems in actual applications of bus travel time prediction. Therefore, we chose to use an SVM to predict bus travel times in this work.

### 3. Methodology

#### 3.1. Prediction Framework

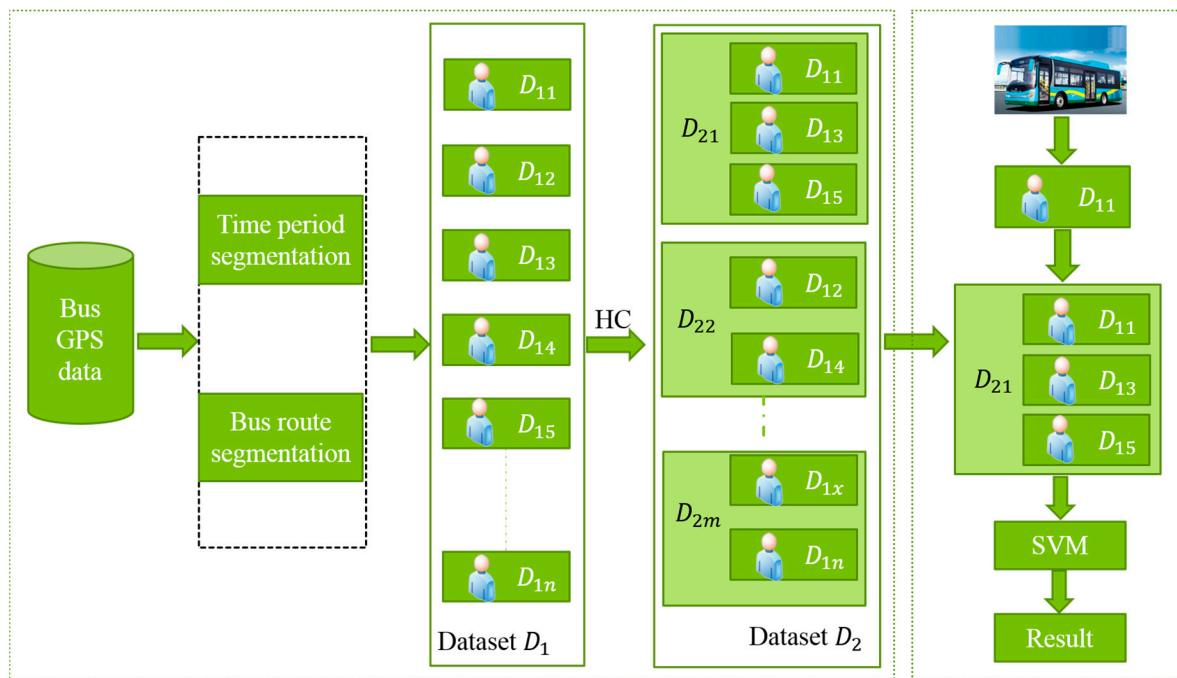
We provide an approach to predicting bus travel times based on similarities in drivers' driving styles. It consists of two main parts: driving style classification and travel time prediction. Figure 2 shows an overview of our proposed approach.

The first part is the hierarchical clustering driver driving style classification model, which is used to classify the driver's driving style according to the bus GPS data. The second part is the bus travel time prediction model. First, the corresponding training dataset was selected according to the classification results of the first part. Then, the corresponding travel time prediction was made for buses driven by drivers with different driving styles. These two parts are described in detail in the following two sections: Sections 3.2 and 3.3.

#### 3.2. Driver Driving Style Classification Method

In order to better distinguish the driving styles of bus drivers, we also consider the characteristics of space and time. First, travel times are categorized. Specifically, a day is divided into three time periods: 7:00–9:00, 9:00–16:00, and 16:00–19:00. The purpose of this time division is to ensure that the travel time of bus drivers has a similar pattern in the same time period, as the time spent on the same road section varies at different time periods [7]. For example, some drivers need to spend 8–10 min driving through a certain road segment during the morning rush hour from 7:00 to 9:00, while they only need 5–8 min during off-peak hours from 9:00 to 16:00 in the same road segment. There are even some drivers who can drive through this section in less than 5 min.





**Figure 2.** An overview of our approach. Dataset  $D_2$  is generated by dataset  $D_1$  after hierarchical clustering. HC—hierarchical cluster; SVM—support vector machine.

Secondly, each bus route is divided into sections. In this step, we consider the road section between two adjacent bus stops as the basic prediction unit. We classify the driving style of the drivers based on the travel time between two adjacent stations (sum of driving time and waiting time). For example, a driver who arrives at a bus stop at a scheduled time can be defined as having a normal driving style. A driver with a mild driving style will arrive later than the scheduled time, while a driver with an aggressive driving style will arrive earlier than the scheduled time. This tendency will gradually become larger with the increase in the distance travelled, so eventually the bus behind will catch up with or overtake the bus in front after a few stops. Similarly, a bus’s waiting time at the bus stop can also reflect the driving style of the driver. Different drivers have different reactions to the speed of passengers’ movements when they get on the bus. For example, if an impatient driver urges passengers to hurry up and get on the bus at the stop, then the stopping time of this type of driver at the stop will be relatively shorter. A gentle driver generally does not urge passengers but waits quietly for passengers to get on the bus calmly, so this type of driver will wait at the stop for a relatively long time. Drivers with a driving style between the two mentioned above will judge whether to urge passengers or not depending on the time. When there is plenty of time, they will not urge passengers, while when time is tight, they will urge passengers to accelerate boarding.

In this work, we use the hierarchical clustering analysis model to classify the driving styles of bus drivers. Hierarchical clustering (HC) is a statistical method used to divide objects into groups with similar meanings. It attempts to find groups that minimize differences within groups and maximize differences between external groups. When we do not know the number of groups in advance but want to create groups and analyze group members, we usually use hierarchical clustering analysis [18]. A hierarchical clustering analysis starts with a single point as a cluster and then continuously merges two clusters until only one cluster is left [37]. This method is described in Algorithm 1.

---

**Algorithm 1:** HC—Basic Hierarchical Clustering Algorithm

---

Input: Sample set  $D = \{x_1, x_2, \dots, x_m\}$ ;  
 Cluster distance metric function  $d$ ;  
 Number of clusters  $k$ .  
 Output: Clusters  $C = \{C_1, C_2, \dots, C_K\}$   
 Process:  
 1: for  $j = 1, 2, \dots, m$  do  
 2:    $C_j = \{x_j\}$   
 3: end for  
 4: for  $i = 1, 2, \dots, m$  do  
 5:   for  $j = i + 1, \dots, m$  do  
 6:      $M(i, j) = d(C_i, C_j)$ ;  
 7:      $M(j, i) = M(i, j)$   
 8:   end for  
 9: end for  
 10: Set the current number of clusters:  $q = m$   
 11: while  $q > k$  do  
 12:   Find the two closest clusters  $C_{i^*}$  and  $C_{j^*}$ ;  
 13:   Merge  $C_{i^*}$  and  $C_{j^*}$ ;  $C_{i^*} = C_{i^*} \cup C_{j^*}$ ;  
 14:   for  $j = j^* + 1, j^* + 2, \dots, q$  do  
 15:     Rename cluster  $C_j$  to  $C_{j-1}$   
 16:   end for  
 17:   Delete the  $j^*$  row and  $j^*$  column of the distance matrix  $M$ ;  
 18:   for  $j = 1, 2, \dots, q - 1$  do  
 19:      $M(i^*, j) = d(C_{i^*}, C_j)$ ;  
 20:      $M(j, i^*) = M(i^*, j)$   
 21:   end for  
 22:    $q = q - 1$   
 23: end while

---

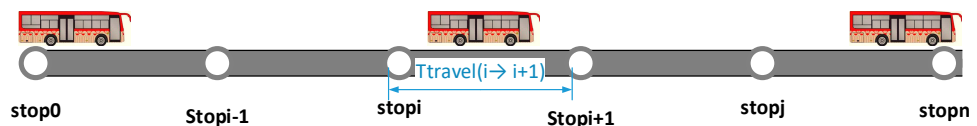
3.3. Bus Travel Time Prediction Method

The objective of predicting bus travel time is to predict bus travel times between two locations (e.g., two bus stops) [3]. The travel time of a bus between two adjacent stops includes the time spent waiting before passing through an intersection, the time spent waiting on the way, and the time spent waiting at the departure stop, as shown in Equation (1).

$$T_{travel(i \rightarrow i+1)} = T_{d,i+1} - T_{d,i} \tag{1}$$

where:

- $T_{travel(i \rightarrow i+1)}$  is the predicted time for the bus to travel from stop  $i$  to stop  $i + 1$ , including the parking time at stop  $i$ , as shown in Figure 3;
- $T_{d,i}$  is the predicted arrival time of the bus at stop  $i$ ;
- $T_{d,i+1}$  is the predicted arrival time of the bus at stop  $i + 1$ ;



**Figure 3.** The time it takes for the bus to travel from stop  $i$  to stop  $i + 1$ .

Similarly, the bus travel time between two non-adjacent stops is the sum of the travel time of multiple adjacent stops.

In this section, we will introduce the process of predicting bus travel time based on clustering drivers’ driving styles. This process includes two steps: (1) the selection of training dataset based on the results of the driver’s driving style classification in the previous

section; and (2) the construction of the travel time prediction model. Corresponding travel time prediction models for drivers with different types of driving styles were constructed.

Previous studies [7,34] have indicated that the kernel of the radial basis function (RBF) is better at predicting bus travel time. Therefore, the RBF kernel function is used in the SVR model in this study, along with the parameters  $C$  and  $e$  of SVR. The selection of this parameter is based on the best combination in previous papers [31], where  $(C, e)$  is set to  $(2, 0.1)$ .  $C$  defines the cost of the penalty function. To accurately and quickly predict the travel time of buses, it is important to determine the appropriate factors for estimating the traffic conditions. The factors we consider in this work are as follows [38]:

- (1)  $X_1$  = day of the week. In general, traffic flow is different during the five working days of the week;
- (2)  $X_2$  = number of road segments. The number of sections between two adjacent stops of the predicted bus line. Different road sections have different road condition characteristics;
- (3)  $X_3$  = departure time of the bus. At different times of the day, the bus journey time is different. For example, during peak hours and off-peak hours, the travel time of buses varies greatly.

The prediction model of bus travel time can be summarized as the following relationship:

$$T_{\text{predicted}} = f(X_1, X_2, X_3) \quad (2)$$

## 4. Case Study

### 4.1. Data Collection and Processing

In this section, experiments were conducted using the Shenyang City 239 bus route as an example to verify the effectiveness of the proposed travel time prediction method described in the last section. The 239 bus route map is shown in Figure 4. The 239 bus line starts from Kang li Automobile Company in the west to the Quan Yuan community in the east. There are 26 stops in the entire journey with a total length of 13.9 km. The operating time period of this bus line is from 06:00 am to 20:30 pm. All buses in the 239 bus line are equipped with GPS positioning devices to collect the location data of each bus every 5 s and transmit them to the data center in real time. The data recorded are listed in Table 1. The serial number of the vehicle-mounted device on each bus is unique and corresponds to a fixed driver. To maintain consistency, only the driving data from the west to the east were used in this study.



Figure 4. The map of bus route no. 239.



**Table 1.** Description of GPS dataset.

Variable	Description
<i>O_LINENAME</i>	the line name
<i>O_TERMINALNO</i>	the serial number of the vehicle-mounted device
<i>O_DATE</i>	the date of data generated
<i>O_TIME</i>	the time of data generated
<i>O_LONGITUDE</i>	Longitude
<i>O_LATITUDE</i>	Latitude
<i>O_SPEED</i>	instantaneous speed
<i>O_UP</i>	running direction
<i>O_NEXTSTATIONNO</i>	the serial number of the next stop

Perhaps due to the influence of GPS signal strength, some road sections have incomplete data. Therefore, the data from the fourth stop to the thirteenth stop were collected and used as the research object in this study (see Figure 4). The corresponding road section consists of nine adjacent stops. For example, road section 7 represents the road section between the sixth stop and the seventh stop, road section 13 represents the road section between the twelfth stop and the thirteenth stop, and so on.

In order to verify the influence of the drivers' driving styles on the travel time, the interference of other factors has to be considered. The time period selected in this study was sunny, and the traffic conditions were normal (no traffic accidents). The buses to be predicted in our model were roughly the same in terms of vehicle type and vehicle performance. Therefore, it can be considered that the main factor affecting the bus travel time was the driving styles of individual drivers.

The data in this study were collected over 11 working days from 4 January 2016 to 18 January 2016. A total of 3204 valid history records of bus travel times were obtained. The data of the first 10 days (4 January 2016 to 15 January 2016) were used as the training data, and the data of the remaining day (18 January 2016) were used as the prediction data.

#### 4.2. Performance Measures

In order to compare the prediction accuracy of the different methods intuitively, two terms, i.e., the mean absolute error (*MAE*) and root mean square error (*RMSE*), are used to analyze the accuracy of our experimental results. Each measure is calculated as follows:

$$MAE = \frac{\sum |t_{observed} - t_{predicted}|}{N} \quad (3)$$

where  $t_{observed}$  is the observed value,  $t_{predicted}$  is the predicted value, and  $N$  is the total number of datasets.

The *MAE* can better reflect the actual situation of the prediction error and also reflect the accuracy of the prediction.

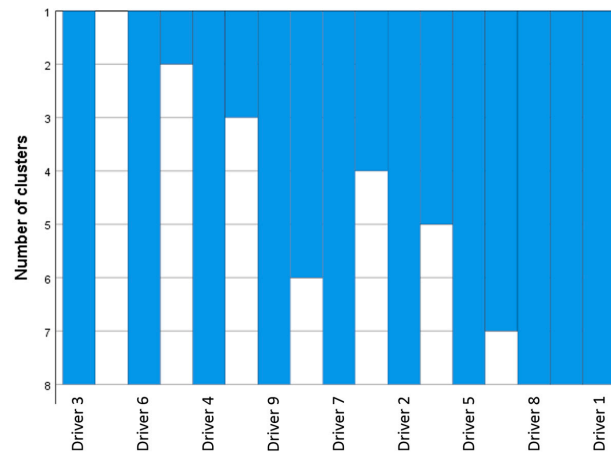
$$RMSE = \sqrt{\frac{\sum (t_{observed} - t_{predicted})^2}{N - 1}} \quad (4)$$

The *RMSE* can express the relative error of the prediction and reflect the stability of the prediction.

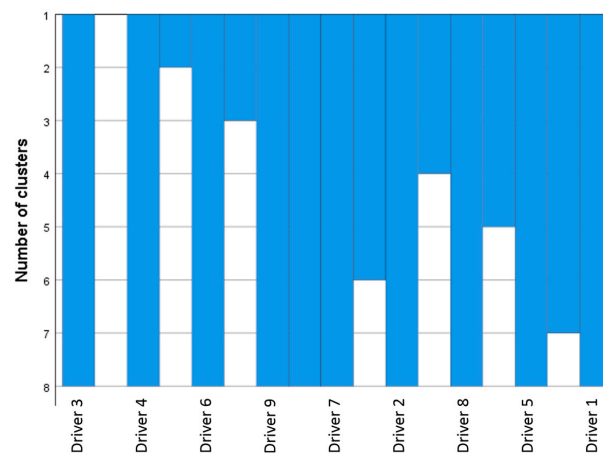
#### 4.3. Driving Style Cluster of Bus Drivers

In this section, the drivers were clustered according to their driving styles. First, a segmentation of the time period was performed. One day was divided into three periods according to the time division scheme proposed in Section 3.2. In each period, the travel time data of nine adjacent stations (5–13) and nine drivers on the no. 239 bus line are stored in the form of a matrix, and then the data processed into the 2D matrix structure are hierarchically clustered.

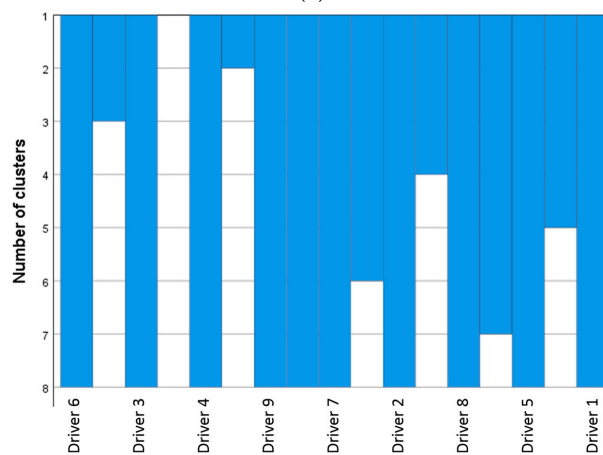
The results are shown in Figure 5. For example, during the morning peak hours from 7:00 to 9:00, when the cluster number is nine, the driving styles are divided into nine categories: {902334, 902355, 902349, 902335}; {902353, 902359}; {902347}; {902351}; and {902340}. In other time periods, the classification results are {902353, 902359, 902335}; {902334, 902349, 902355}; {902347}; {902349}; and {902355}.



(a)



(b)



(c)

**Figure 5.** Hierarchical clustering results of three time periods. The corresponding numbers of drivers 1, 2, 3, 4, 5, 6, 7, 8, and 9 are 902334, 902335, 902340, 902347, 902349, 902351, 902353, 902355, and 902359. (a) Morning peak time: 7:00–9:00. (b) Off-peak time: 9:00–16:00. (c) Afternoon peak time: 16:00–19:00.

#### 4.4. Results

In order to select the best dataset for the prediction of travel time, four different levels of datasets are selected from the clustering results in the previous section of the experiment, which are divided into nine categories, five categories, four categories, and one category. For example, the training datasets corresponding to the four prediction models of driver number 902334 are shown in Table 2. Among them, Model 1 is when the dataset is clustered into nine categories, that is, the travel time history data of a single driver are used for training and prediction. Model 4 is trained and predicted by the travel time history data of all drivers when the dataset is clustered into one category. Model 2 and Model 3 use the historical data of 3–6 drivers’ travel time for training and prediction when they are clustered into five and four categories, respectively, according to the clustering results.

**Table 2.** The training datasets corresponding to the four prediction models of driver number 902334.

Model	Number of Clusters	7:00–9:00	9:00–16:00	16:00–19:00
Model 1	9	902334	902334	902334
Model 2	5	902334, 902335, 902349, 902355	902334, 902349, 902355	902334, 902349, 902359
Model 3	4	902334, 902335, 902349, 902353, 902355, 902359	902334, 902335, 902349, 902353, 902355, 902359	902334, 902335, 902349, 902353, 902355, 902359
Model 4	1	All data	All data	All data

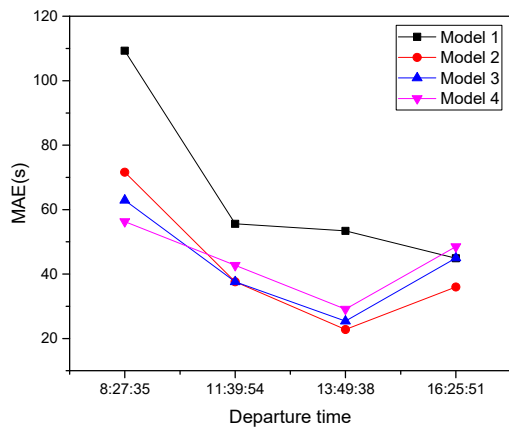
The four models were experimentally evaluated, and six drivers were selected as experimental subjects on 18 January 2016. Table 3 lists the scheduled departure schedules of the six predicted drivers on 18 January 2016. The travel time of each driver will be predicted using four models from different datasets.

**Table 3.** The schedule of the six predicted drivers on 18 January 2016.

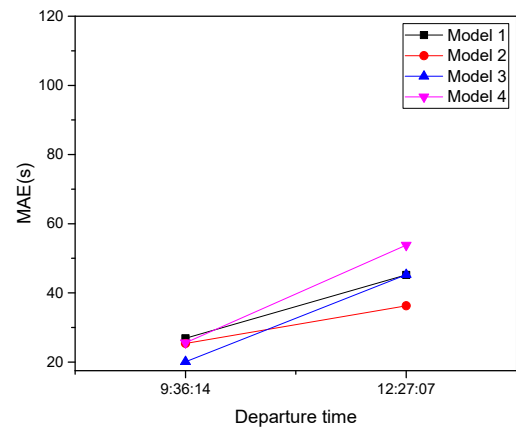
Time Period	902334	902335	902351	902353	902355	902359
7–9	8:27:35		7:37:39	7:16:42		7:21:04
9–12	11:39:54	9:36:14	11:26:30	10:25:45	9:47:19	10:14:51
12–16	13:49:38	12:27:07	13:54:07	14:19:43	12:39:59 15:22:28	15:13:35
16–19	16:25:51		16:40:31	17:05:54		17:45:58

Figure 6 shows the MAE values of the predicted results of all the trips of the buses driven by the six drivers under the four prediction models. Table 4 is the average MAE of all the trips of each bus in one day under the four prediction models that are used for the forecast. The global MAE (*g-MAE*) score measures the average MAE of all the buses driven by the six drivers in the four models. The bold results are the results with the best performance for this experiment. (The bold results in the following tables have the same meaning.)

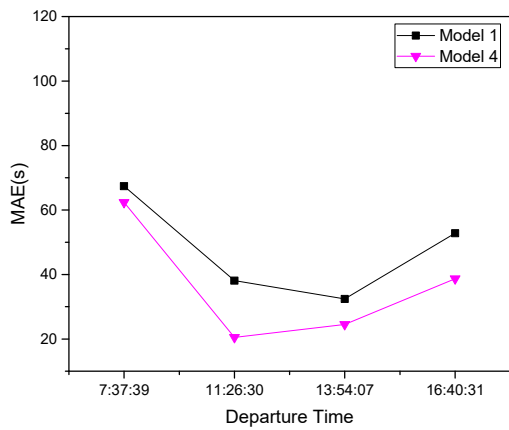
The results show that the performance of Model 2 and Model 3 is better than that of Model 1 using only a single driver history dataset and Model 4 using all the datasets, and the prediction error of Model 2 is the smallest (compared to model 1, the accuracy is improved by 24.5%, compared to Model 3, the accuracy is improved by 9.1%, and compared to Model 4, the accuracy is improved by 13.4%). Obviously, the prediction effect of Model 1 using only the historical data of a single driver is the most unsatisfactory, which may be the result of too few training data. However, the accuracy of the prediction results is not as simple as the more data, the better. The prediction effect of Model 4 using all the data is not the best. We think that the main reason for this is that the driving behaviors of the drivers are different, resulting in a wide distribution of travel time data.



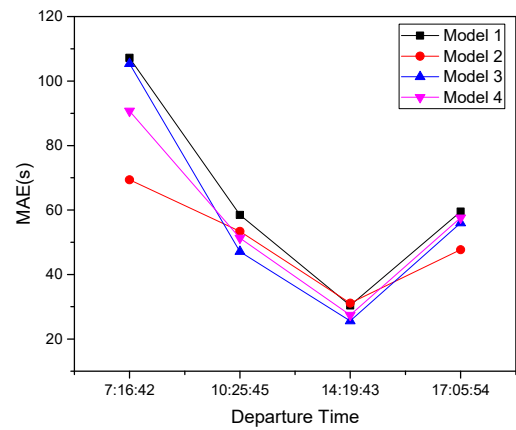
(a)



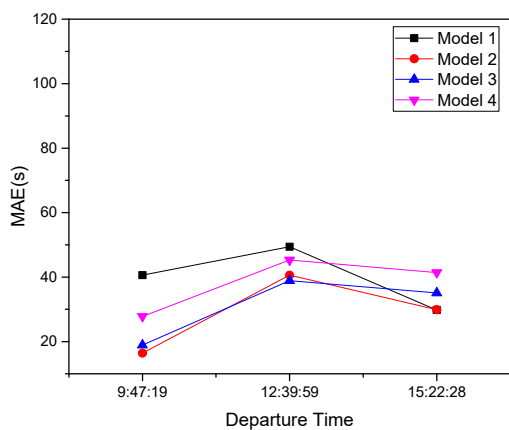
(b)



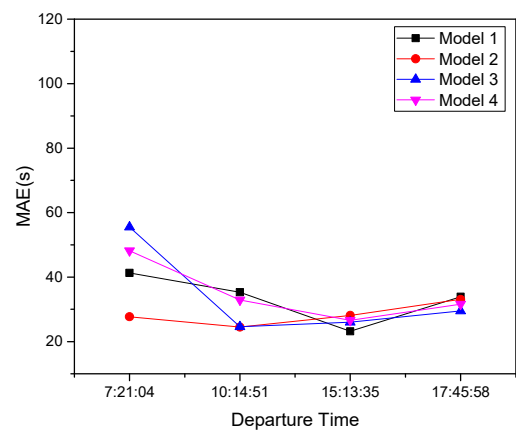
(c)



(d)



(e)



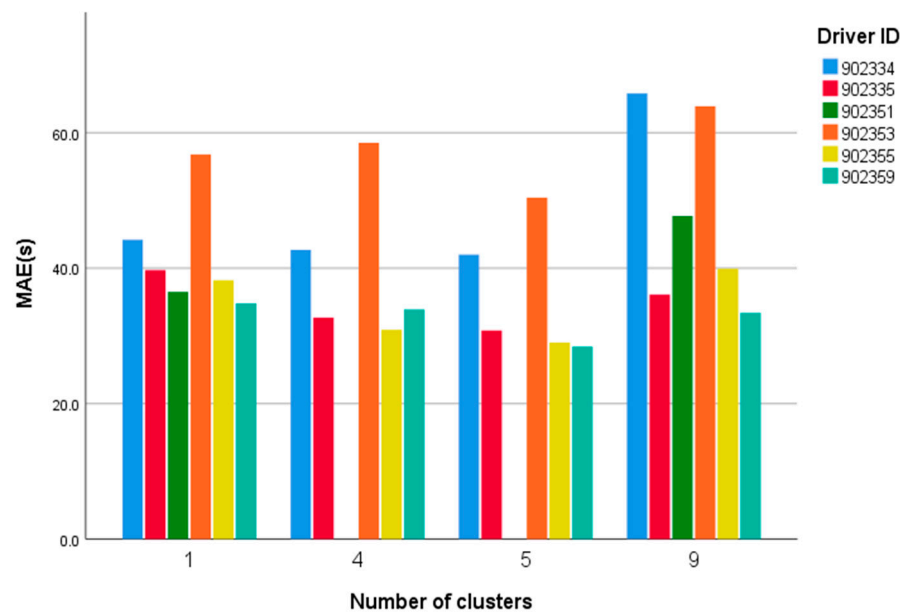
(f)

**Figure 6.** The MAE values of the predicted results of all trips of the buses driven by six drivers under the four prediction models. (a) 902334. (b) 902335. (c) 902351. (d) 902353. (e) 902355. (f) 902359.

**Table 4.** Prediction results of four models (*MAE*).

Driver	Model 1	Model 2	Model 3	Model 4
902334	65.8	42.0	42.7	44.2
902335	36.1	30.8	32.7	39.7
902351	47.7			36.5
902353	63.9	50.4	58.5	56.8
902355	39.9	29.0	30.9	38.2
902359	33.4	28.4	33.9	34.8
<i>g-MAE</i>	47.8	36.1	39.7	41.7

Figure 7 shows the impact of different numbers of clusters on the performance of the model. As the number of clusters increases, the *MAE* value first decreases and then increases, indicating that a smaller number of clusters is not enough to distinguish different driving style data. A larger dataset cannot provide personalized information for the model. On the other hand, a larger number of clusters can lead to a smaller amount of data in some categories, resulting in the poor predictive performance of the model. This indicates that selecting appropriate predictive data is crucial.

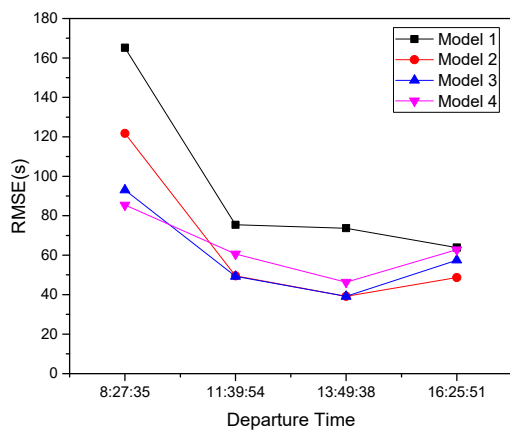


**Figure 7.** The Influence of different clustering results on model prediction performance (*MAE*).

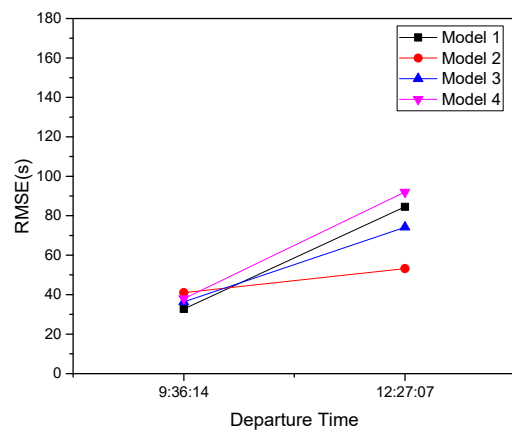
Figure 8 shows the comparison results of the root mean square error (*RMSE*) of the four prediction models, and Table 5 shows the average *RMSE* results of all the trips of each predicted bus the next day using the four prediction models. The global *RMSE* (*g-RMSE*) score measures the average *RMSE* of the buses driven by the six drivers in the four models.

The results show that the performance of Model 2 and Model 3 is better than that of Model 1 using only a single driver history dataset and Model 4 using all datasets, and the *RMSE* value of Model 2 is the best (24.6% higher than that of Model 1, 10.2% higher than that of Model 3, and 15.6% higher than that of Model 4).

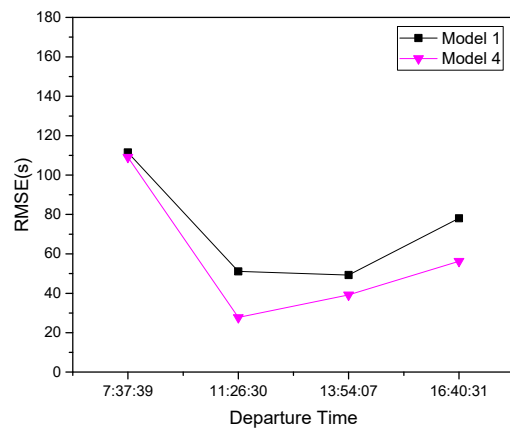




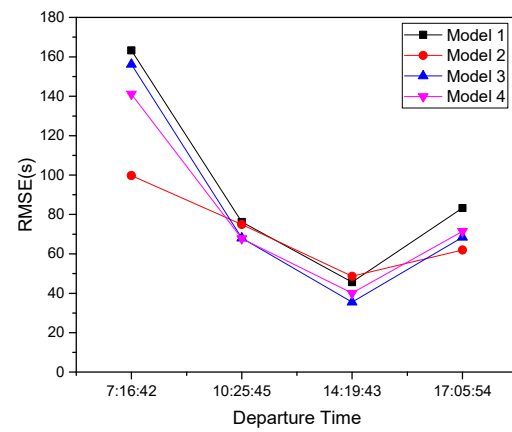
(a)



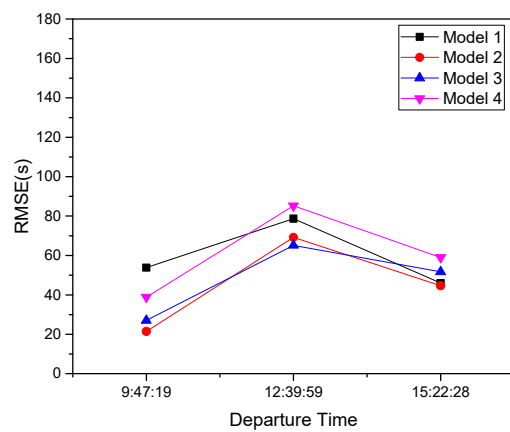
(b)



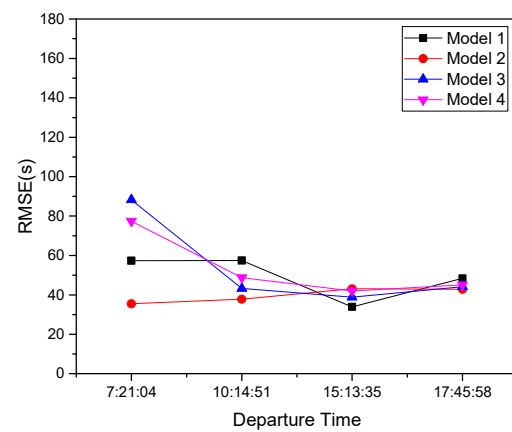
(c)



(d)



(e)



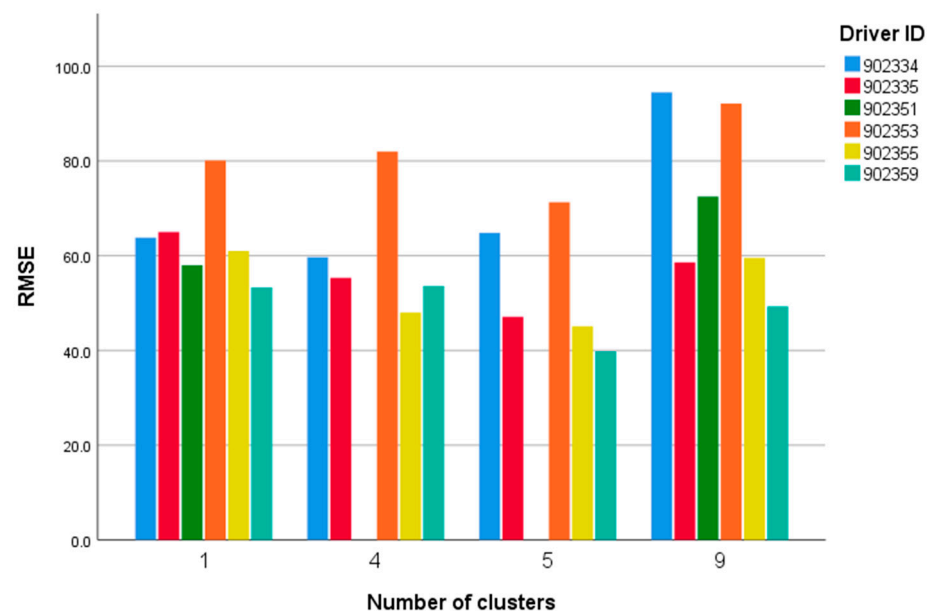
(f)

**Figure 8.** The RMSE values of the predicted results of all trips of the buses driven by six drivers under the four prediction models. (a) 902334. (b) 902335. (c) 902351. (d) 902353. (e) 902355. (f) 902359.

**Table 5.** Prediction results of four models (*RMSE*).

Driver	Model 1	Model 2	Model 3	Model 4
902334	94.5	64.8	59.7	63.8
902335	58.6	47.1	55.3	65.0
902351	72.5			58.0
902353	92.1	71.3	82.0	80.1
902355	59.5	45.1	48.0	61.0
902359	49.3	39.8	53.6	53.3
<i>g-RMSE</i>	71.1	53.6	59.7	63.5

The same conclusion can be drawn from the results in Figure 9: as the number of clusters increases, the *RMSE* value also decreases at first and then increases. This indicates that both too large and too small datasets are not conducive to model prediction.

**Figure 9.** The influence of different clustering results on model prediction performance (*RMSE*).

#### 4.5. Discussion

In this study, a new method combining the hierarchical clustering algorithm and support vector regression model was proposed to study the influence of driving styles on the prediction of travel time. In previous studies on the prediction of bus travel times, other scholars have considered the impact of changes in passenger flow [39], traffic conditions [40,41], space–time factors [42–44], signals [45–47], weather [40], and other factors on the prediction of travel time. Compared to these influencing factors, it is more difficult to obtain driving style data. Therefore, to our knowledge, there has not been research that has taken into account driving styles in the prediction of bus travel times. In addition, although there are a lot of traffic data being collected currently, they lack accurate processing. How to extract feature sets that can accurately predict bus travel times from these data is still worth studying [45]. Recently, some meaningful research has been conducted on this subject. For example, buses' running times for multiple routes were used to predict arrival times for each route, which improves prediction precision [30]. Ma et al. [44] uses clustering based on the homogeneity of travel time observations and potential traffic conditions to estimate the probability distribution of travel time according to the travel time distribution of road sections. He et al. [40] explored the common travel time patterns of different bus line sections and divided bus line sections with

similar patterns into the same cluster. Then these clusters were merged to extract data records for model training and the prediction of bus travel times. This paper also considers driving styles from the perspective of optimizing the prediction dataset. For the first time, driving styles were used to optimize the training and prediction dataset, and then the optimized datasets were used to predict the bus travel times. The effectiveness of the proposed method is verified by experiments.

Moreover, from our experimental results, we also found that the *MAE* values of the four models were higher than those of other periods in the early peak hours. One of the possible reasons for this is that the route selected for research in this paper is one-way, driving from the outside of the city to the center of the city. The number of passengers and vehicles entering the city in the early hours of the morning is greater than in other periods, leading to more uncertain factors in terms of travel time. We believe that the impact of traffic conditions and drivers' driving styles on bus travel times is both relevant and restrictive. Subsequent research should consider the comprehensive impact of multiple factors, which can further improve the accuracy of predictions. Although the prediction quality is low during peak hours, our framework can also significantly improve prediction accuracy.

## 5. Conclusions

In this work, a method for predicting bus travel times by clustering drivers' driving styles was presented. The historical travel times of different bus drivers were first clustered, and then the prediction of the travel time was made using the historical dataset of the corresponding driver, which achieved the objective of improving the speed and precision of the travel time prediction. Our experiments were conducted using a large-scale dataset of real-world data collected in Shenyang, China. The experimental results indicate that the bus travel time prediction method (Model 2) based on the clustering of drivers' driving styles proposed in this research paper increases the average accuracy by 13.4%. In addition, it is found that the performance of Model 2 and Model 3 (the models that are based on the clustering results after the dataset selection) is better than that of Model 1 (the model using only a single driver history dataset) and Model 4 (the traditional method using all the datasets). Among all the models, Model 2 has the best prediction accuracy, which is in contrast to Model 1, which has the worst prediction accuracy. We believe that the main reason for this difference is the sparsity of the dataset of individual drivers. At the same time, it is also found that the prediction accuracy of Model 4 is not the best, indicating that it is not correct to assume that the more data, the better the accuracy of the prediction results. We believe that the main reason for this is that using data from all drivers can mask differences in driver behaviors.

In our future work, we will consider more behavioral factors, such as bus drivers' acceleration behavior, braking behavior, and overtaking behavior, to further study the impact of bus drivers' driving styles on travel time.

**Author Contributions:** All authors contributed to the study's conception, the design of the experiments, and the paper's structure. Z.Y. performed the experiment analysis and wrote the first draft of the manuscript. All authors participated in the revision and proofreading of the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Key Project of National Natural Science Foundation of China (U1908212), the Central Government's Guided Local Science and Technology Development Fund Project (1653137155953), and Liaoning Province's "Takes the Lead" Science and Technology Research Project (2021jh1/10400006).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Jeong, R.; Rilett, L.R. Bus arrival time prediction using artificial neural network model. In Proceedings of the 7th IEEE Intelligent Transportation System Conference, Washington, DC, USA, 3–6 September 2004. [\[CrossRef\]](#)
2. Fan, W.; Gurmu, Z. Dynamic Travel Time Prediction Models for Buses Using Only GPS Data. *Int. J. Transp. Sci. Technol.* **2015**, *4*, 353–366. [\[CrossRef\]](#)
3. Yu, B.; Wang, H.; Shan, W.; Yao, B. Prediction of bus travel time using random forests based on near neighbors. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 333–350. [\[CrossRef\]](#)
4. Wei, M.; Liu, X. How wet is too wet? Modelling the influence of weather condition on urban transit ridership. *Travel Behav. Soc.* **2022**, *27*, 117–127. [\[CrossRef\]](#)
5. O’Sullivan, A.; Pereira, F.C.; Zhao, J.; Koutsopoulos, H.N. Uncertainty in Bus Arrival Time Predictions: Treating Heteroscedasticity with a Metamodel Approach. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 3286–3296. [\[CrossRef\]](#)
6. Yin, T.; Zhong, G.; Zhang, J.; He, S.L.; Ran, B. A prediction model of bus arrival time at stops with multi-routes. *Transp. Res. Procedia* **2017**, *25*, 4627–4640. [\[CrossRef\]](#)
7. Ma, J.; Chan, J.; Ristanoski, G.; Rajasegarar, S.; Leckie, C. Bus travel time prediction with real-time traffic information. *Transp. Res. Part C Emerg. Technol.* **2019**, *105*, 536–549. [\[CrossRef\]](#)
8. Daganzo, C.F. A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. *Transp. Res. Part B Methodol.* **2009**, *43*, 913–921. [\[CrossRef\]](#)
9. Cats, O. Regularity-driven bus operation: Principles, implementation and business models. *Transp. Policy* **2014**, *36*, 223–230. [\[CrossRef\]](#)
10. Yao, B.; Hu, P.; Lu, X.; Gao, J.; Zhang, M. Transit network design based on travel time reliability. *Transp. Res. Part C Emerg. Technol.* **2014**, *43*, 233–248. [\[CrossRef\]](#)
11. Yu, H.; Chen, D.; Wu, Z.; Ma, X.; Wang, Y. Headway-based bus bunching prediction using transit smart card data. *Transp. Res. Part C Emerg. Technol.* **2016**, *72*, 45–59. [\[CrossRef\]](#)
12. Martinez, C.M.; Heucke, M.; Wang, F.-Y.; Gao, B.; Cao, D. Driving Style Recognition for Intelligent Vehicle Control and Advanced Driver Assistance: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 666–676. [\[CrossRef\]](#)
13. Ahn, K.; Rakha, H.; Trani, A.; Van Aerde, M. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. *J. Transp. Eng.* **2002**, *128*, 182–190. [\[CrossRef\]](#)
14. Mudgal, A.; Hallmark, S.; Carriquiry, A.; Gkritza, K. Driving behavior at a roundabout: A hierarchical Bayesian regression analysis. *Transp. Res. Part D Transp. Environ.* **2014**, *26*, 20–26. [\[CrossRef\]](#)
15. Johnson, D.A.; Trivedi, M.M. Driving style recognition using a smartphone as a sensor platform. In Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), Washington, DC, USA, 5–7 October 2011; pp. 1609–1615. [\[CrossRef\]](#)
16. Murphey, Y.L.; Milton, R.; Kiliaris, L. Driver’s style classification using jerk analysis. In Proceedings of the 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems, Nashville, TN, USA, 30 March–2 April 2009; pp. 23–28. [\[CrossRef\]](#)
17. Langari, R.; Won, J.-S. Intelligent energy management agent for a parallel hybrid vehicle—Part I: System architecture and design of the driving situation identification process. *IEEE Trans. Veh. Technol.* **2005**, *54*, 925–934. [\[CrossRef\]](#)
18. Doshi, A.; Trivedi, M.M. Examining the impact of driving style on the predictability and responsiveness of the driver: Real-world and simulator analysis. In Proceedings of the 2010 IEEE Intelligent Vehicles Symposium, La Jolla, CA, USA, 21–24 June 2010; pp. 232–237. [\[CrossRef\]](#)
19. Miyajima, C.; Nishiwaki, Y.; Ozawa, K.; Wakita, T.; Itou, K.; Takeda, K.; Itakura, F. Driver modeling based on driving behavior and its evaluation in driver identification. *Proc. IEEE* **2007**, *95*, 427–437. [\[CrossRef\]](#)
20. Constantinescu, Z.; Marinoiu, C.; Vladoiu, M. Driving style analysis using data mining techniques. *Int. J. Comput. Commun. Control* **2010**, *5*, 654–663. [\[CrossRef\]](#)
21. Vaitkus, V.; Lengvenis, P.; Zylis, G. Driving style classification using long-term accelerometer information. In Proceedings of the 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR), Miedzyzdroje, Poland, 2–5 September 2014; pp. 641–644. [\[CrossRef\]](#)
22. Karginova, N.; Byttner, S.; Svensson, M. *Data-Driven Methods for Classification of Driving Styles in Buses*; SAE International: Warrendale, PA, USA, 2012. [\[CrossRef\]](#)
23. Augustynowicz, A. Preliminary classification of driving style with objective rank method. *Int. J. Automot. Technol.* **2009**, *10*, 607–610. [\[CrossRef\]](#)
24. Xu, L.; Hu, J.; Jiang, H.; Meng, W. Establishing style-oriented driver models by imitating human driving behaviors. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2522–2530. [\[CrossRef\]](#)
25. Zhu, L.; Yu, F.R.; Wang, Y.; Ning, B.; Tang, T. Big data analytics in intelligent transportation systems: A survey. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 383–398. [\[CrossRef\]](#)
26. Patnaik, J.; Chien, S.; Bladikas, A. Estimation of bus arrival times using APC data. *J. Publ. Transp.* **2004**, *7*, 1–20. [\[CrossRef\]](#)
27. Chien, S.I.-J.; Ding, Y.; Wei, C. Dynamic bus arrival time prediction with artificial neural networks. *J. Transp. Eng.* **2002**, *128*, 429–438. [\[CrossRef\]](#)
28. Chen, M.; Liu, X.; Xia, J.; Chien, S.I. A dynamic bus-arrival time prediction model based on APC data. *Comput.-Aided Civ. Infrastruct. Eng.* **2004**, *19*, 364–376. [\[CrossRef\]](#)

29. van Lint, J.W.C.; Hoogendoorn, S.P.; van Zuylen, H.J. Accurate freeway travel time prediction with state-space neural networks under missing data. *Transp. Res. Part C Emerg. Technol.* **2005**, *13*, 347–369. [[CrossRef](#)]
30. Yu, B.; Lam, W.H.K.; Tam, M.L. Bus arrival time prediction at bus stop with multiple routes. *Transp. Res. Part C Emerg. Technol.* **2011**, *19*, 1157–1170. [[CrossRef](#)]
31. Wu, C.-H.; Ho, J.-M.; Lee, D.T. Travel-time prediction with support vector regression. *IEEE Trans. Intell. Transp. Syst.* **2004**, *5*, 276–281. [[CrossRef](#)]
32. Vanajakshi, L.; Rilett, L.R. Support vector machine technique for the short term prediction of travel time. In Proceedings of the 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, 13–15 June 2007; pp. 600–605. [[CrossRef](#)]
33. Yu, B.; Yang, Z.Z.; Yao, B.Z. Bus arrival time prediction using support vector machines. *J. Intell. Trans. Syst.* **2006**, *10*, 151–158. [[CrossRef](#)]
34. Shalaby, A.; Farhan, A. Prediction models of bus arrival and departure times using AVL and APC data. *J. Public Transp.* **2004**, *7*, 41–61. [[CrossRef](#)]
35. Vanajakshi, L.; Subramanian, S.C.; Sivanandan, R. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. *IET Intell. Transp. Syst.* **2009**, *3*, 1–9. [[CrossRef](#)]
36. Bai, C.; Peng, Z.-R.; Lu, Q.-C.; Sun, J. Dynamic Bus Travel Time Prediction Models on Road with Multiple Bus Routes. *Comput. Intell. Neurosci.* **2015**, *2015*, 432389. [[CrossRef](#)]
37. Zhou, Z.H. *Machine Learning*, 1st ed.; Tsinghua University Press: Beijing, China, 2016; pp. 214–216.
38. Yin, Z.; Zhang, B. Construction of Personalized Bus Travel Time Prediction Intervals Based on Hierarchical Clustering and the Bootstrap Method. *Electronics* **2023**, *12*, 1917. [[CrossRef](#)]
39. Shalaby, A.; Farhan, A. Bus travel time prediction model for dynamic operations control and passenger information systems. In Proceedings of the 82nd Annual Meeting of the Transportation Research Board, Washington, DC, USA, 12–16 January 2003.
40. He, P.; Jiang, G.; Lam, S.-K.; Sun, Y. Learning heterogeneous traffic patterns for travel time prediction of bus journeys. *Inf. Sci.* **2020**, *512*, 1394–1406. [[CrossRef](#)]
41. Huang, Y.P.; Chen, C.; Su, Z.C.; Chen, T.S.; Sumalee, A.; Pan, T.L.; Zhong, R.X. Bus arrival time prediction and reliability analysis: An experimental comparison of functional data analysis and Bayesian support vector regression. *Appl. Soft Comput.* **2021**, *111*, 107663. [[CrossRef](#)]
42. Kumar, B.A.; Vanajakshi, L.; Subramanian, S.C. Bus travel time prediction using a time-space discretization approach. *Transp. Res. Part C Emerg. Technol.* **2017**, *79*, 308–332. [[CrossRef](#)]
43. Hua, X.; Wang, W.; Wang, Y.; Ren, M. Bus arrival time prediction using mixed multi-route arrival time data at previous stop. *Transport* **2017**, *33*, 543–554. [[CrossRef](#)]
44. Ma, Z.; Koutsopoulos, H.N.; Ferreira, L.; Mesbah, M. Estimation of trip travel time distribution using a generalized Markov chain approach. *Transp. Res. Part C Emerg. Technol.* **2017**, *74*, 1–21. [[CrossRef](#)]
45. Yuan, Y.; Shao, C.; Cao, Z.; He, Z.; Zhu, C.; Wang, Y.; Jang, V. Bus Dynamic Travel Time Prediction: Using a Deep Feature Extraction Framework Based on RNN and DNN. *Electronics* **2020**, *9*, 1876. [[CrossRef](#)]
46. Bie, Y.; Wang, D.; Qi, H. Prediction Model of Bus Arrival Time at Signalized Intersection Using GPS Data. *J. Transp. Eng.* **2012**, *138*, 12–20. [[CrossRef](#)]
47. Chow, A.H.F.; Li, S.; Zhong, R. Multi-objective optimal control formulations for bus service reliability with traffic signals. *Transp. Res. Part B Methodol.* **2017**, *103*, 248–268. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.