Retraction

Retracted: Signal Parameters Estimation and Optimization Using Mobile Navigation Data

Mathematical Problems in Engineering

Mathematical Problems in Engineering has retracted the article titled "Signal Parameters Estimation and Optimization Using Mobile Navigation Data" [1], due to concerns that the submitting author, Luxi Dong, did not obtain permission from the coauthors prior to the submission and publication of the article. Luxi Dong has clarified that they did not make a significant contribution to the study and the study was completed by Jiyuan Tan and the other coauthors. All listed authors agree to the retraction of this article for the above reasons.

References

[1] L. Dong, X. Liu, M. Jia, J. Tan, and X. Wei, "Signal Parameters Estimation and Optimization using Mobile Navigation Data," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8898938, 15 pages, 2020.

Research Article

Signal Parameters Estimation and Optimization Using Mobile Navigation Data

Luxi Dong ,¹ Xiaoming Liu ,¹ Meijie Jia ,² Jiyuan Tan ,¹ and Xiangda Wei ¹

¹*Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Beijing 100144, China*
²*Capital Aerospace Machinery Corporation Limited, Beijing 100076, China*

Correspondence should be addressed to Luxi Dong; 281667195@qq.com

In the management and evaluation of traffic network, signal parameters are important for monitoring and evaluating the operation state and the traffic capacity of intersection. However, a wide range of real-time signal timing schemes lacks a clear and effective method. In this paper, we propose the signal parameter calculation method based on mobile navigation data. Then, the possibility of crossing intersection passing time of the stop line is studied. The time differences between passing times of different cycles are distributed periodically that several peaks appear cycle by cycle. The relationship between sampling rate and relative error is discussed. Combined with the distribution peak normality test, the appropriate distribution peak is selected through the actual case. The cycle lengths and effective red time parameters are calculated and compared with the known signal parameters. The result demonstrates the proposed method has high accuracy and provides data support for the research of the traffic management.

1. Introduction

Why do we choose mobile navigation data to estimate the signal timing? First of all, the timing parameters of most intersections cannot be obtained directly, which need to be acquired by the system. Due to the permission, time is not so easy to obtain. Another method is investigated at the intersection, but it will consume a huge number of human resources. However, the popularity of mobile navigation data is provided us with the possibility of estimating timing parameters. The advantages of high precision and coverage of mobile navigation data make it possible to estimate the time alignment; it is simple and effective to use these data [1].

Many scholars put forward a variety of methods for the estimation of signal timing parameters. Hao et al. [2] proposed a method to estimate the timing of the signal based on single lane and no overtaking. Combined with the traffic flow theory and learning optimization method, a three-step method of estimating the timing parameters is proposed, including cycle breaking estimation, exact cycle boundary detection, and effective red (green) time estimation. The

method of the start and end of a period and the effective green and red time is estimated based on the delayed model [3]. According to continuous time series of traffic flow for signal timing estimation in Ban's research, it needs to observe the effective trajectory of continuous cycles. However, most of the research required high data sampling rate; the existing domestic floating car data or mobile navigation data could not meet the data sampling rate demands of such methods [4].

Although the basic idea is the effective trajectory of the intersection based on public mobile navigation data, the time is not taken into account [5]. Ren et al. [6] provided fluctuation similarity measure, such as dynamic time warping and gray relation grade, and the hierarchical clustering algorithm was used to further separate the traffic flow time series. Jeff Ban et al. [7] presented a model that requires sampled travel times between two consecutive positions on main roads, one upstream and the other downstream of a signalized intersection, without need to know the signal timing or traffic flow information. The model proposed two observations regarding delays for

signalized intersections: (a) delay approximately showed piecewise linear curves due to the characteristics of queue forming and dissipating; (b) there was obvious increasing in delay after the start of the red time that enables detection of the start of a cycle. The model and algorithm were verified by reasonable results based on experiment data. In order to match travel time characteristics of the vehicle at two locations, Kwong et al. [8] constructed a statistical model without measurements of signal settings. The signal settings could be inferred from the matched vehicle results. Kerper et al. [9] provided the Traffic Light Coordination Analysis (TLCorA) to calculate from traffic light whether there was a representative approaching trajectory. A dynamic time warping algorithm was applied to classify the approaching trajectory. Fayazi et al. [10] demonstrated the feasibility of estimating traffic signal phase and timing from statistical patterns based on low-frequency vehicular probe data. Their method was reduced empty time at red signals and improved fuel efficiency and lower emissions. Zhao et al. [11] presented an improved car-following model accounting for the driver's characteristics and automation for longitudinal driving. Stability analysis is performed for both driver's characteristics and controller gains adopting frequency domain sweeping method. Zhao et al. [12] proposed Optimal Transmission Reliability Enhancement Mechanism (OTREM) for the development of the cooperative driving systems; it can integrate the vehicular cyber system with the vehicular physical system for the optimization of the cooperative driving at traffic intersections. Tong et al. [13] proposed a stochastic programming (SP) model to schedule adaptive signal timing plans that minimize the expected vehicle delay in oversaturated state. The results show that SP model was better than the deterministic linear programming (LP) model in total vehicle delay. Li et al. [14] defined and determined the potential dependence among time series data. Then, a decomposition algorithm was used to separate daily-similar trend and nonstationary bursts components from the traffic flow time series based on the Granger test. The findings revealed the relationship between the structure of road networks and the correlations among traffic time series. At the same time, Li et al. [15] constructed a long-term and short-term trend model of traffic time series. The proposed model could improve prediction accuracy and not only specified the temporal pattern but also related it to the spatial relation of traffic time series. Axer et al. [16] studied the periodicity of the vehicle trajectory in the fixed signal timing and then estimated the cycle start time by calculating the time difference between the reference time and the real trajectory time stamp. Moreover, Axer and Friedrich [17] proposed a method that calculated the stage of red-light duration; it took the trajectory to pass through the stop line and estimate a possible cycle length based on time module. All of these results are useful for different saturations. Fayazi et al. [18] extracted an estimated collection of signal phase and timing (SPaT) information based on real-time feed of sparse and low-frequency probe vehicle data. The results could be applied in the field of safety driving assistant. Wang et al. [19] detected effectively real curved trajectories occurring at traffic intersections. The heterogeneity of traffic density was considered when using the curved trajectories to automatically infer the actual cycle.

In summary, the phase sequence is determined based on the mobile navigation data in this paper. Firstly, the clustering method is used to estimate the cycle length based on Tan's red-light estimated model. Secondly, the second derivative is applied to explore the mutation point to estimate the red time. Finally, the good results are obtained.

The rest of this paper is arranged as follows: Section 2 introduces the description vehicles data used in this paper. Section 3 provides the clustering method to calculate the cycle length and effective red time parameters. The result is shown in Section 4. Finally, the conclusions are given in Section 5.

2. Data Description

2.1. Time Distribution and Spatial Distribution.
There are currently about three types of vehicle driving data: one is provided by Traffic Committee, the other is collected by car-hailing service such as DIDI, Uber, and so on, and the rest is navigation data. The navigation data of three intersections in Beijing is used, accumulated historical data collected from January to July 2019. There are about 145496 vehicle trajectories per intersection per day, proving a good coverage of three intersections. The sampling rate of the data is generally collected about 5 s (see Table 1), which mainly contains vehicle id, date, time, latitude, longitude, vehicle speed, and road id (see Table 2). The vehicle trajectories are obtained based on latitude and longitude.

2.2. Road Matching.
The MapInfo Professional is used to complete the road matching. Each road matches a corresponding id in Figure 1; for example, the two section ids of the east entrance at the intersection are represented, respectively, by 22473 and 27704. If any trajectory data's road id is the same as one of the two ids, it will be identified as on the road. Thus, this method is used to handle all the data points, which is matched to the actual road network.

2.3. Abnormal Data.
In the process of data preprocessing, some abnormal trajectories are encountered, where the time distribution, distance, and direction are irrational.

Time error is defined as follows: the vehicle driving recording time is within the appointed time range. Moreover, the downstream recording time is greater than the upstream recording time.

Direction error is defined as follows: when the time is satisfied, it is determined whether the vehicle's driving direction number and its change are consistent with the searched change of valid trajectory driving direction number.

Distance error is defined as follows: after the above requirements are satisfied, it is determined whether the distance between the starting point and the ending point of the trajectory satisfied the selected distance.

Reentry error means that it is also necessary to determine whether such a trajectory returned itself during the traveling.

TABLE 1: The proportion of time interval.

Interval (s)	5	10	12	11	1	6	2	4
Proportion (%)	76.16	2.93	2.84	2.01	1.81	1.62	1.38	1.37

TABLE 2: The description of data.

Vehicle id	Date	Time	Latitude	Longitude	Vehicle speed	Road id
1	2019-02-18	08:10:26	39.99238200	116.40754700	21	54919
1	2019-02-18	08:10:31	39.99527700	116.40775300	20	54893
. . .	. . .	. . .	. . .	. . .	. . .	. . .
1	2019-02-18	22:15:20	39.99464800	116.40759300	46	48395

FIGURE 1: Road in MapInfo professional.

2.4. Parameter Calculation. With the analysis of mobile navigation data, the following parameters are calculated as follows: travel time, delay, and arriving time at the stop line. The travel time and delay are indices that can be evaluated as traffic conditions at an intersection. At the same time, the travel time and delay are closely related to the red time. These parameters are used to estimate the red time. However, the time passing at the stop line of the same phase of the vehicles also shows a certain regularity.

2.4.1. Arriving Time at the Stop Line. As shown in Figure 2, this is a vehicle trajectory that the direction is west to east. The upstream and downstream lines are observation lines, index_1 and index_2 are the nearest points beside the upstream line, and index_3 and index_4 are the nearest points beside the downstream line. t_1, t_2, t_3, and t_4 are the time of the four points. L_1, L_2, L_3, and L_4 are the distances of index points and observation lines. Assume that the car is driven in uniform motion during the short distance, so the time is allocated based on distance.

$$t_{\text{downstream}} = t_3 + \frac{t_4 - t_3}{L_3 + L_4} L_3, \tag{1}$$

$$t_{\text{upstream}} = t_1 + \frac{t_2 - t_1}{L_2 + L_1} L_1. \tag{2}$$

2.4.2. Travel Time and Delay. From the above formulas, the time of vehicles passing the observation lines and stop lines is calculated. The travel time of a vehicle passing intersections can be obtained through the difference between $t_{\text{downstream}}$ and t_{upstream}.

$$t_{\text{traveltime}_i} = t_{\text{downstream}} - t_{\text{upstream}}. \tag{3}$$

Delay is an index of traffic conditions at an intersection. If the travel time of a vehicle is known, then the delay time at the intersection is estimated:

$$d_{\text{Delay}_i} = t_{\text{traveltime}_i} - \frac{l}{v_{\text{free}}}, \tag{4}$$

where l is the distance of the vehicle passing the intersection and v_{free} is the free flow speed.

So, we can calculate the travel time, delay, and the moment of passing the stop line of each vehicle through the intersection base on the mobile navigation data, with the abnormal data removed.

3. Method

The red time and cycle length are two important parameters of signal timing which enable the following two models. The first model uses single-phase delay and red-light duration to estimate the red time, and then multi-phase delay is used to
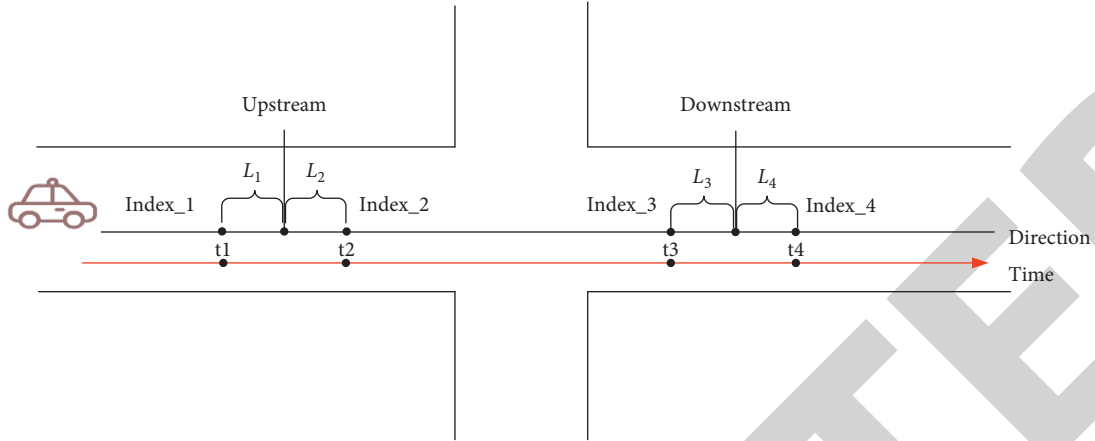
FIGURE 2: Passing time calculation.

estimate the cycle length in [5]. The second model uses difference distribution of single-phase passing time to estimate the red light and cycle time. The flowchart of method is shown in Figure 3.

3.1. Cycle Length Estimation.

The vehicles passing the intersections are controlled and influenced by the signal timing. The periodic distribution of time passing through the stop line is like a simple two-stage intersection. To make the traffic flow go through intersection more safely and non-stopped, it is of important to analyze cycle length estimation. An example is organized as follows.

The traveling vehicles encounter a red time before the stop line. The vehicles should be stopped and queued one by one. When the traffic lights turn green, the vehicles will be started to another queue and leave the intersection in turn. Thus, there are vehicles that directly pass the stop line or stay before the stop line waiting for the red time end. During the peak hours in the city, the vehicles are usually jammed and stopped by traffic lights. In addition, the vehicles cannot completely dissipate in front of the stop line during one red time; it is necessary to require two or more red times to pass the stop line.

The number of vehicles passing through the stop line during the north-south phase green light in several signal cycles is analyzed. As shown in Figure 4, the red lines present the vehicles passing stop line during the red time of north and south phases, which is green time of the other phases. time_difference_1 is the time difference of passing time during green time and waiting for one red time, and time_difference_2 is the time difference of passing time during green time and waiting for two red times.

As shown in Figure 5, if car2 and car3 can be extracted, then t_{red} is calculated as follows:

$$t_{\mathrm{car3}} - t_{\mathrm{car2}} = \text{time\_difference} = t_{\mathrm{red}}, \tag{5}$$

where car3 is the first car that passes the stop line during green time, t_{car2} is the latest car that passes the stop line during green time of the last cycle, and t_{red} is the effective red time.

Due to the sampling rate, car2 and car3 are not real sampled; there may be car1, car4, car5, car6, or any car among them. Thus, the data is sampled maybe from car1 to car2 and car3 to car4. If the data is large enough, it will present a peak distribution. The first peak range is $[r, r + 2g]$. The second peak of the distribution is $[2r + g, 2r + 3g]$, while the interval between two adjacent peaks is one cycle length.

The difference between the mean values of every two adjacent distribution peaks $(r + g)$ is estimated as cycle length. Therefore, the cluster distribution method is used to calculate the mean of the distribution peaks. The peak distribution presents four or five in the normal circumstance, which can prove the data sampling rate is appropriate. Otherwise, the following conditions will happen. When the sampling rate is high, each selected period is divided. Thus, the differences are calculated as follows:

$$C = \frac{\sum_{i=1}^{N} \text{mean}(i+1) - \text{mean}(i)}{N}, \tag{6}$$

where N is the number of peaks. As shown in Figure 6, the different numbers of distribution peaks appear under different data sizes. Different numbers of distribution peaks lead to different results. More detailed relationships are discussed in Section 4.

Table 3 shows the cycle length estimation of four phases at 06:35 to 12:00. In most of the results, the relative errors are within an acceptable range. However, the relative error of north to south is 33%. There are only two peaks in this phase. The reason is that the number of trajectories is too much so that the range of differences is smaller than the other. The relationship between the number of peaks and sampling rate is discussed in Section 4.

3.2. Effective Red Time Estimation.

In the ideal case, the minimum value of the time_difference distribution peak is an effective red time. However, there is much interference data before the first peak. The key to the method is effectively removing the interference data. The minimum value of the first peak is obtained in Figure 7. The rising gradient method
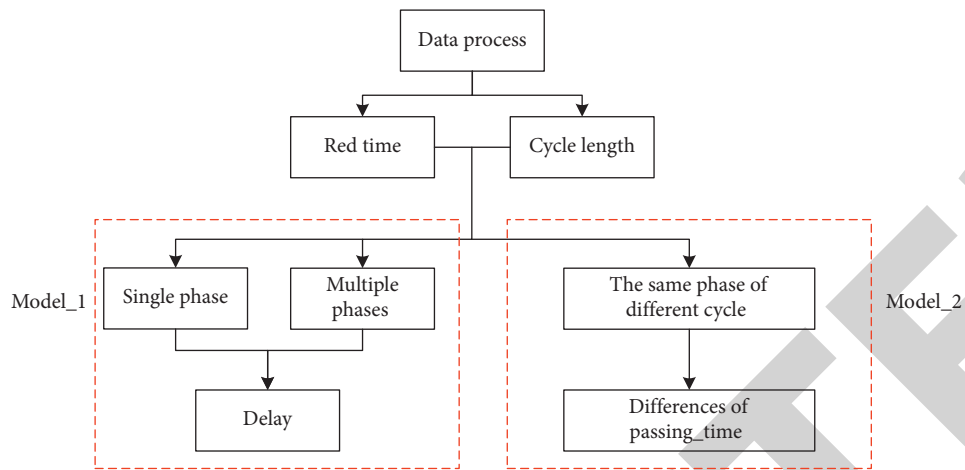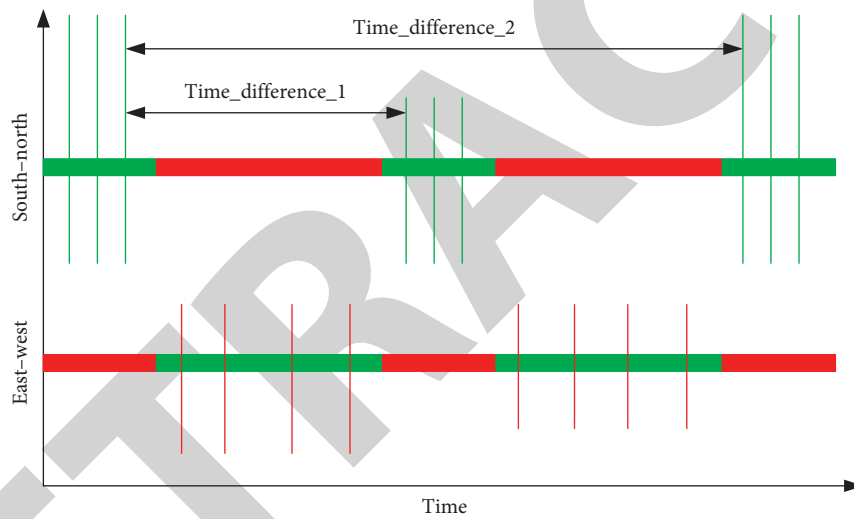
Figure 3: The flowchart of method.

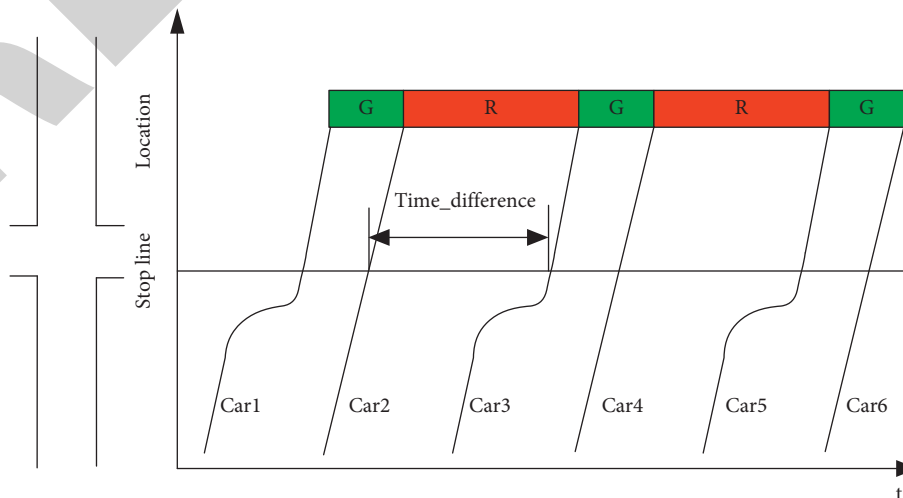Figure 4: Time passing the stop line.

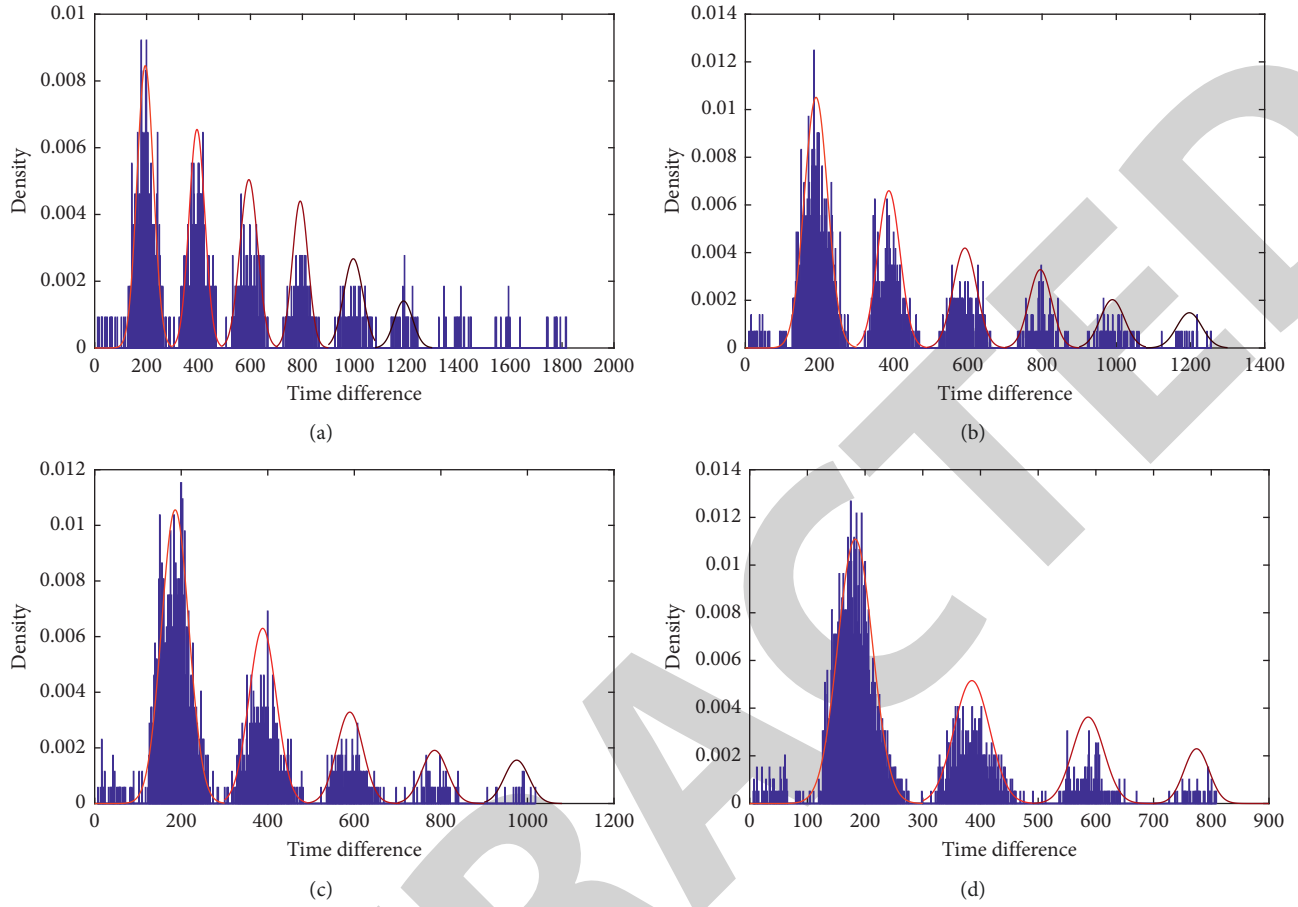Figure 5: The trajectories of passing the stop line.

(a)

(b)

(c)

(d)

Figure 6: The different number of distribution peaks. (a) The number of distribution peaks is 6. (b) The number of distribution peaks is 6 in another case. (c) The number of distribution peaks is 5. (d) The number of distribution peaks is 4.

Table 3: The cycle length estimation.

Phase (s)	Period (s)	Mean 1	Mean 2	Mean 3	Mean 4	Estimation (s)	Real cycle length (s)	Relative error (%)
East to west	06:35–12:00	191	395	596	800	203	215	5.6
West to east	06:35–12:00	191	396	597	797	202	215	6.0
South to north	06:35–12:00	186	389			203	215	5.4
North to south	06:35–12:00	194	338			144	215	33

of the time of vehicle passing the stop line is proposed in [5]. Figure 8 shows the probability density distribution. The red point is the minimum value of the density distribution.

In this paper, another method is proposed to obtain the effective minimum value of time_difference. The empirical distribution functions are used to explore the empirical distribution of each time_difference. The data before the first peak is evenly distributed. When it belongs to the normal distribution, the value of empirical density obviously rises rapidly to produce a catastrophe point (see Figure 8). The second derivative method obtains a catastrophe point. The maximum value corresponding to the time_difference is the effective red time. The catastrophe points are calculated as follows:

$$y_n' = \frac{\Delta y_n}{\Delta x_n} = \frac{y_{n+1} - y_n}{x_{n+1} - x_n}, \tag{7}$$

$$y_n'' = \frac{\Delta y_n'}{\Delta x_n'} = \frac{y_{n+1}' - y_n'}{x_{n+1} - x_n}. \tag{8}$$

As shown in Figure 9, the first peak ranges from about 0 to 300. Figure 9 shows the empirical distribution of time_difference. Figure 10 shows the first derivative of time_difference; the whole curve first increases and then decreases. Then, the first derivative is derived to obtain the second derivative (see Figure 11). The time corresponding to the maximum value of the second derivative is the effective red time.
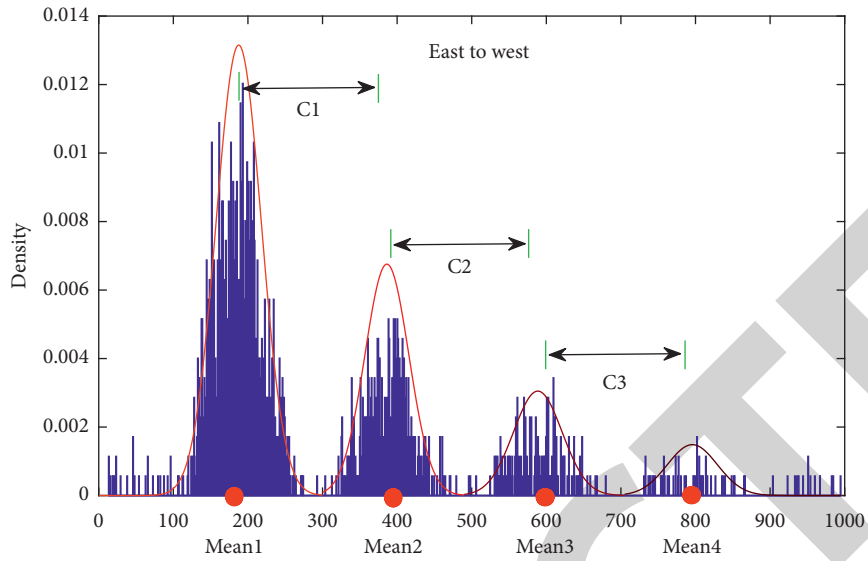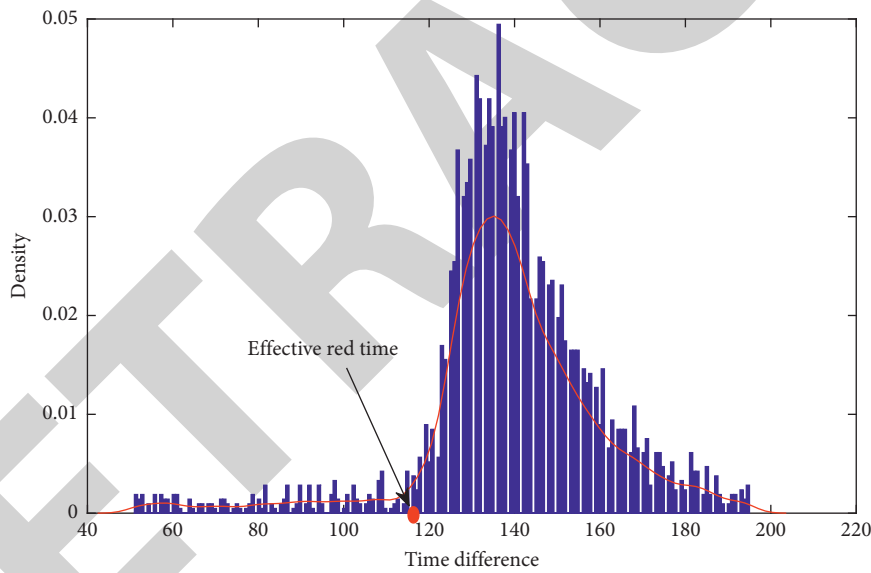
Figure 7: The cycle length estimation.

Figure 8: The probability density distribution.

The red time estimation of four phases at 06:35 to 12:00 is shown in Table 4. All errors of east to west and west to east are slightly larger. It is speculated that the amount of data from south to north or north to south is relatively large; the estimation result could be more accurate. If the sampling rate is higher, the relationship between the relative error and sampling rate is discussed in Section 4.

3.3. Estimation Algorithm Steps. The cycle length and effective red time are estimated effectively based on the above method; the algorithm is divided into several steps.

Step1. The vehicle trajectory data passing the intersection is sampled, and then road matching is completed.

Step2. The time passing stop line is calculated by (1) and (2) based on searching points.

Step3. The time headway of the signal period is calculated by (5) based on the vehicle passing the stop line in two adjacent signal periods. time_difference is the time difference between two or more signal periods.

Step4. All the time differences are summarized, and then the frequency histogram and probability density distribution map are plotted based on time differences.

Step5. The cycle length is calculated by (6).

Step6. The data of first peak is deleted and empirical distribution curve is drawn. The first derivative and second derivative are calculated by (7) and (8), and then the maximum of the second derivative is obtained. The
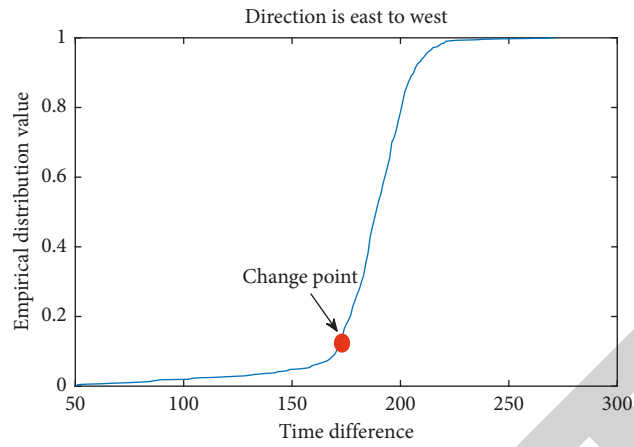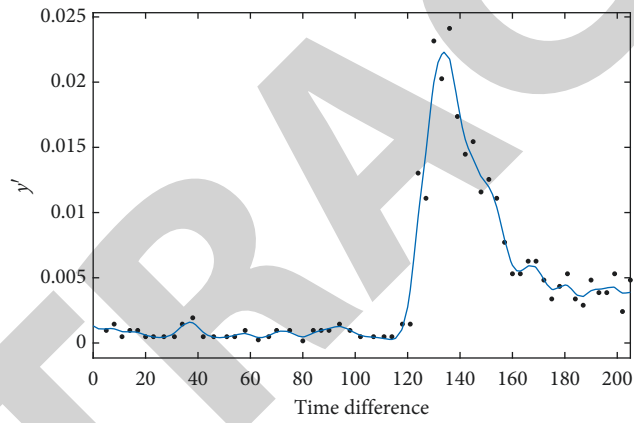
Figure 9: The empirical distribution.

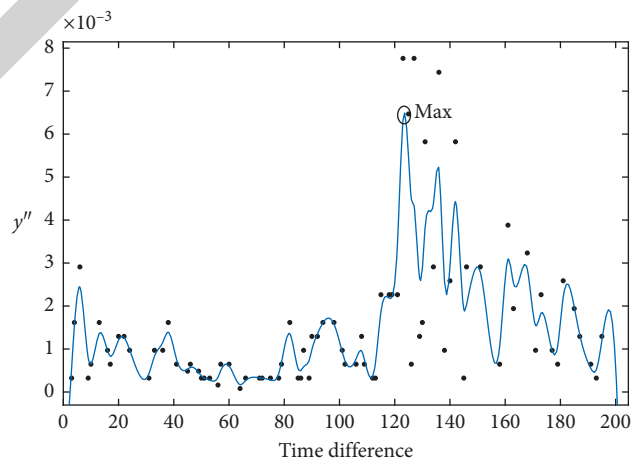Figure 10: The distribution of the first derivative.

Figure 11: The distribution of the second derivative.

Table 4: The red time estimation.

Phase (s)	Period (s)	Red time	Estimation (s)	Relative error (%)
West to east	06:35–12:00	181	170	6.0
North to south	06:35–12:00	119	122	2.5
East to west	06:35–12:00	181	170	6.0
South to north	06:35–12:00	131	127	3.1

time_difference corresponding to this point is estimated effective red time.

4. The Results and Discussion

Although the cycle length and effective red time are estimated based on the above method, there are still some details in the estimation of the cycle length. For example, the number of distribution peaks causes a deviation, and it is not verified whether the distribution peaks belong to normal distributions. The effective distribution peak experiment and the relationship between sampling rate and deviation are discussed as follows.

4.1. The Effective Distribution Peak Experiment. The normal distribution is used to fit the data for each distribution peak. However, it is not verified whether the data conform to normal distribution. Thus, the normal distribution probability graph is combined with the lillietest function, whether each peak satisfies the normal distribution.

Figure 12 presents nearly eight peaks, but only the forward four data peaks satisfy normal distribution. The principle of lillietest is that we assume the data satisfy the normal distribution, and then calculate the parameter h. If h is 0, it means the data satisfy normal distribution. Otherwise, the assumption is not valid. The effective distribution peak algorithm is arranged as follows:

Step 1. The number of data for the peak is calculated, which is judged as whether it is greater than the threshold h (50); if it is not satisfied, then the judgement is ended.

Step 2. The parameter h is calculated by lillietest function. If h is 0, the data satisfies the normal distribution. Otherwise, the assumption is not valid.

Step 3. The normal distribution probability graph is drawn.

Step 4. The data that satisfies normal distribution is fitted.

As shown in Figure 13 and Table 5, the first four peaks are represented as the normal distribution (see Figure 13). The parameter h of the first four peaks is 0 (see Table 5). Thus, the data of the first four peaks are used and fitted.

4.2. The Relationship between Penetration and Deviation. There is only one distribution peak of time_difference in Figure 14; it could not estimate the cycle. The distribution peak of time_difference should be observed.

The sampling rate of the vehicles passing the stop line is changed during green time; we choose 50, 100, 150, 200, 250, and 300 vehicles to test. The statistics of each sampled data and the deviations of appearing peak are shown in Table 6.

n/c is the average vehicle number of each cycle and deviations is the difference between real cycle length and estimation; then we fit the n/c and deviations with the small RMSE in Figure 15. The relationship is fitted as follows:

$$\text{deviation} = 54.42\left(\sin\left(\frac{n}{c} - \pi\right)\right) + \left(-2.44 * \left(\left(\frac{n}{c} - 10\right)^2\right)\right) + 258, \tag{9}$$

where $(n/c) \in [0, 1]$.

There are 0.55 vehicles in each cycle in Figure 15. The number of effective distribution peaks is 4, which has the smallest deviation. The cycle length is able to be estimated accurately. If we want to estimate the cycle length, the n/c is less than 1. Otherwise, it is not an effective distribution peak at all.

According to the n/c after discussion and optimization, the data is rescreened and processed. Compared with the previous results, the north to south phase is changed from two distribution peaks to four distribution peaks. The relative error is decreased from 33% to 7.1% in Table 7. The experimental result of n/c is verified as reasonable.

4.3. The Results. The vehicle trajectories of three intersections are used to verify the method effectiveness. Each intersection is divided into eight phases including four straight phases and four left turn phases. As the method mentioned, we choose a suitable number of peaks of every histogram of time_difference. The mean difference of two adjacent peaks is regarded as one cycle length. The final cycle length is extracted based on the mean of three cycle lengths.

As shown in Figure 16(a), the estimated deviation of cycle length is around 1s in Lincui road and Kehui road intersection. The relative error between the real cycle length and estimation of cycle length is 0∼0.70%. The average of the relative error is 0.38%. The relative error of straight phase is a little higher than left turn phase. The estimated and real cycle length are almost the same. In general, the estimated error at the second intersection is smaller than the first intersection. Thereby, the results indicate that the cycle length estimation method is applied to signals with either fixed or variable cycle length. In Figure 16(b), the relative error between the real cycle length and estimated cycle length is 5.55∼6.80% in Anli road and Huizhong road intersection. The average of the relative error is 5.98%. The relative error of left turn phase is a little higher than straight phase. In Figure 16(c),
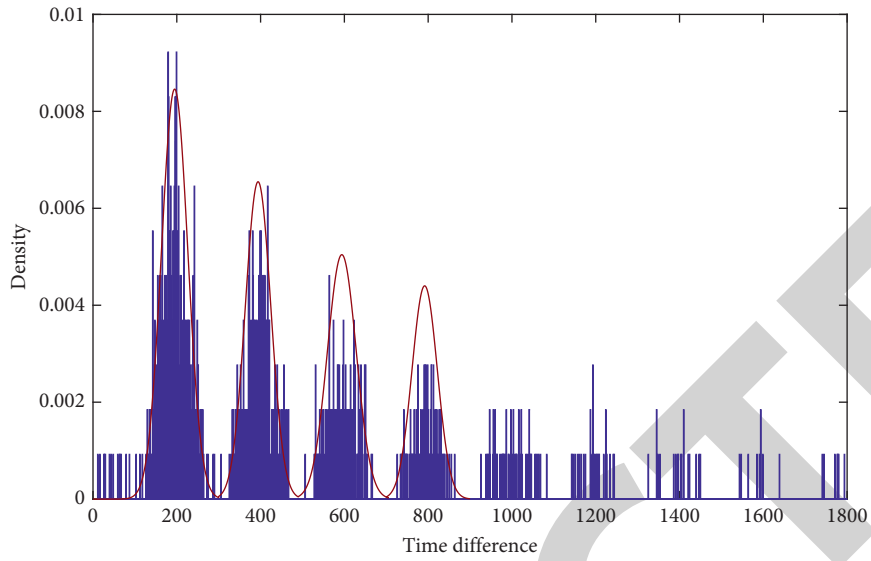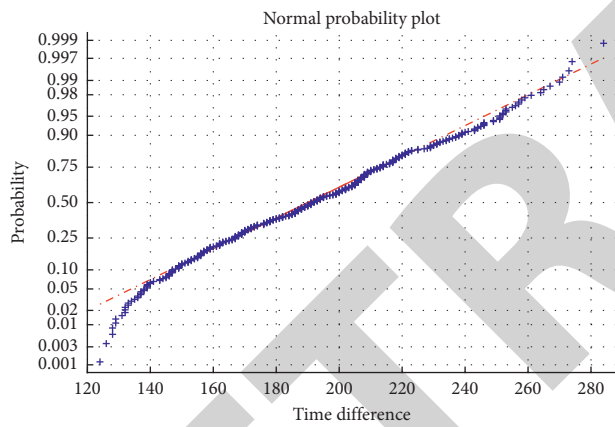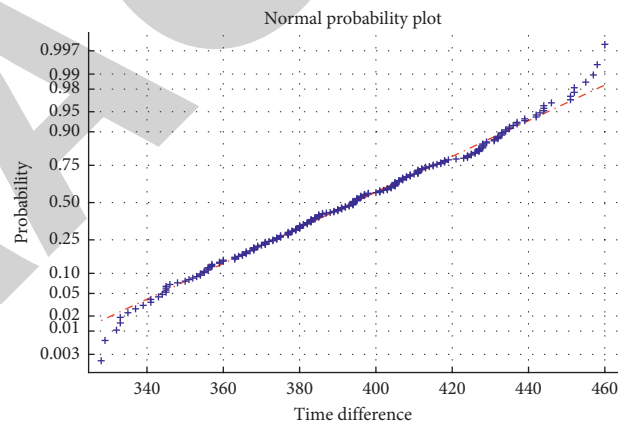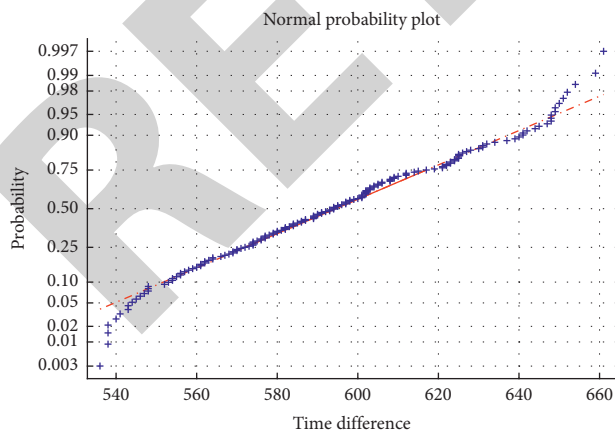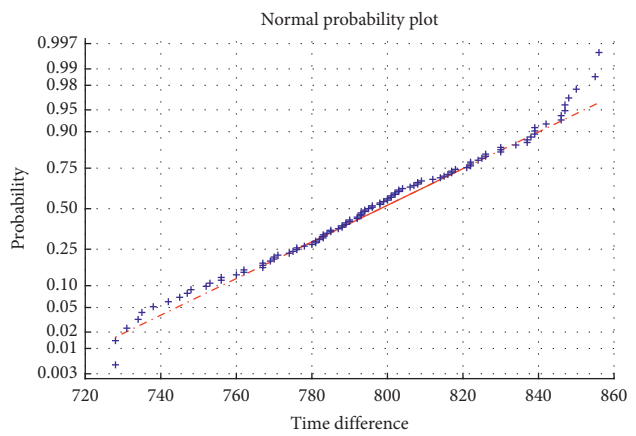
FIGURE 12: The distribution of probability density distribution.

(a)

(b)

(c)

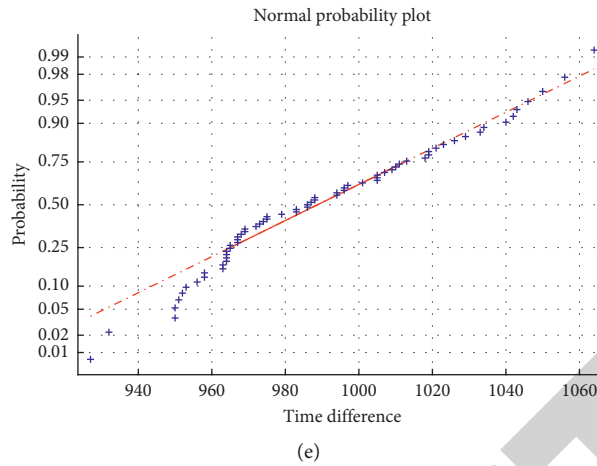(d)

FIGURE 13: Continued.

(e)

Figure 13: The normal probability plot of peaks. (a) Peak 1. (b) Peak 2. (c) Peak 3. (d) Peak 4. (e) Peak 5.

Table 5: The h of each peak.

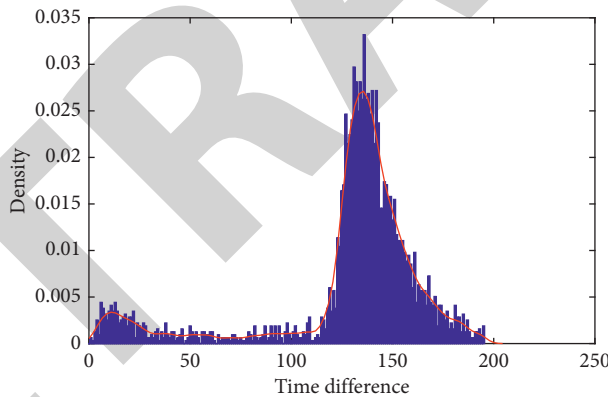| Peak (s)	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5
h	0	0	0	0	1

Figure 14: The probability density distribution.

Table 6: The n/c and deviations.

n/c	0.25	0.5	0.75	1
Deviations	12.44	11.69	11.96	14.5

the relative error between the real cycle length and estimated cycle length is 5.62~6.42% in Anli road and Huizhong north road intersection. The average of the relative error is 5.97%.

Figure 17(a) shows the red time estimation of three intersections. Figure 17(a) presents that most of the relative errors of Lincui road and Kehui road are under 5%. The relative error between the real red time and estimation of red time is 1.58~3.91%. The average of the relative error is 2.43%. In general, the relative error of left turn phase is a little higher than straight phase. As shown in Figure 17(b), the relative error of Anli road and Huizhong road is 4.05% to 6.15%. The average of relative error is 5.21%. Moreover, most of them are under 5%. The Anli-Huizhong intersection is divided into two periods a day. In Figure 17(c), the relative error of Anli road and Huizhong North road is 5.03% to 6.02%. The average of relative error is 5.11%.

Whether it is the estimation of the cycle length or the estimation of the red time, the relative errors of Anli-
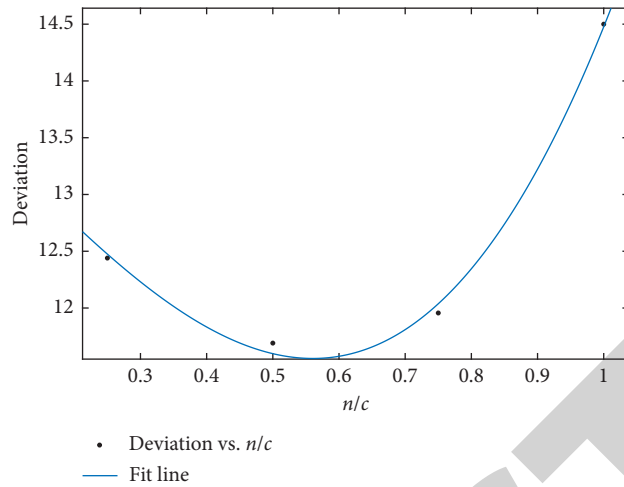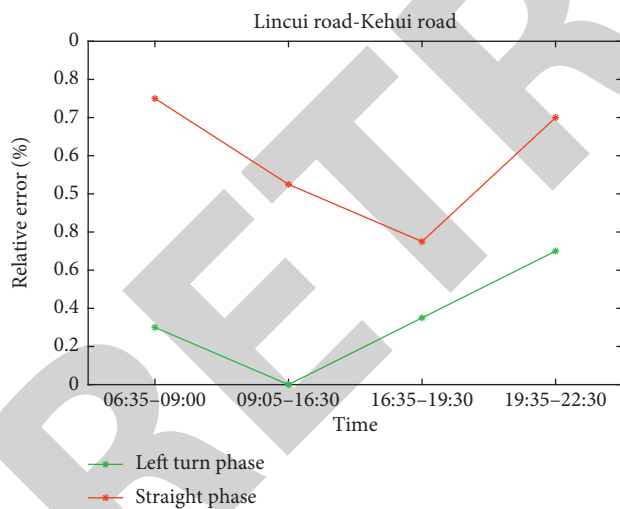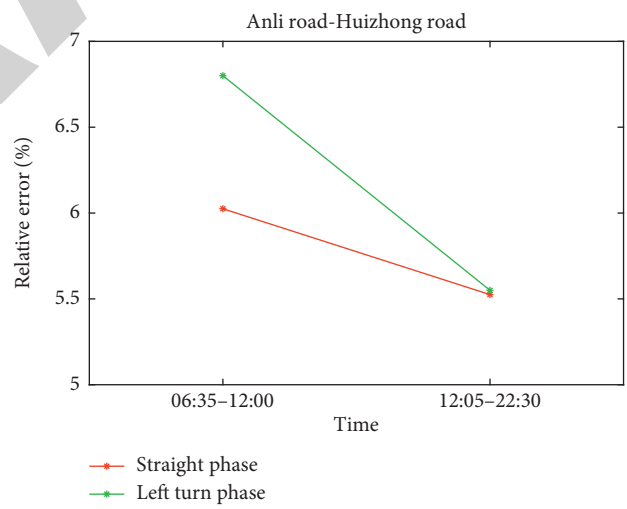
Figure 15: The relationship of n/c and deviations.

Table 7: The cycle length estimation.

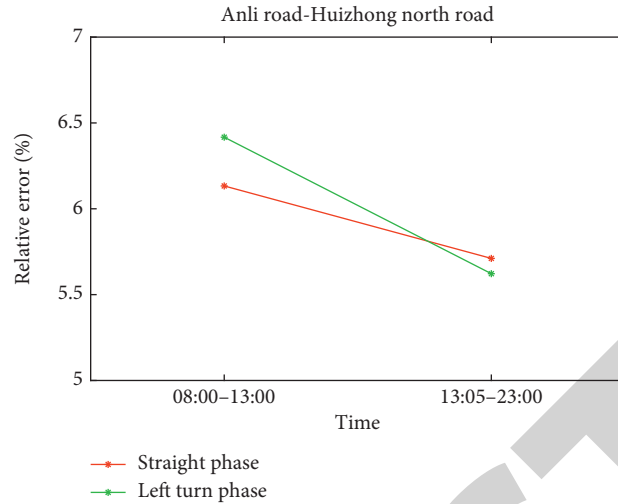| Phase (s)	Period (s)	Mean 1	Mean 2	Mean 3	Mean 4	Estimation (s)	Real cycle length (s)	Relative error (%)
East to west	06:35–12:00	191	395	596	800	203	215	5.6
West to east	06:35–12:00	191	396	597	797	202	215	6.0
South to north	06:35–12:00	186	389	598	796	203	215	5.4
North to south	06:35–12:00	194	338	593	793	200	215	7.1

(a)

(b)

Figure 16: Continued.

(c)

FIGURE 16: (a) The cycle length estimation in Lincui road-Kehui road. (b) The cycle length estimation in Anli road-Huizhong road. (c) The cycle length estimation in Anli road-Huizhong North road.
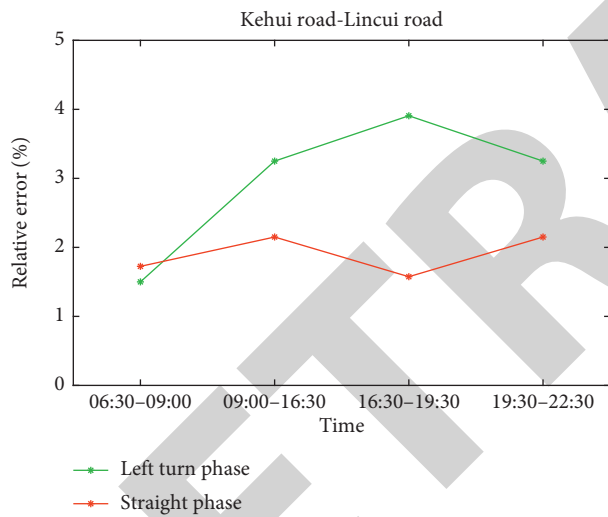
(a)
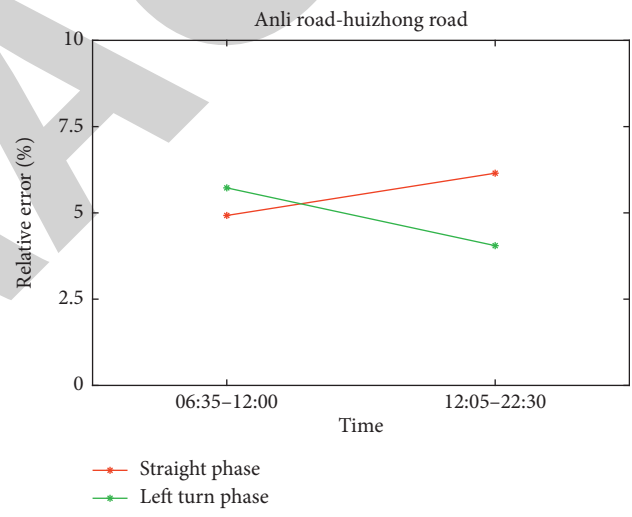
(b)
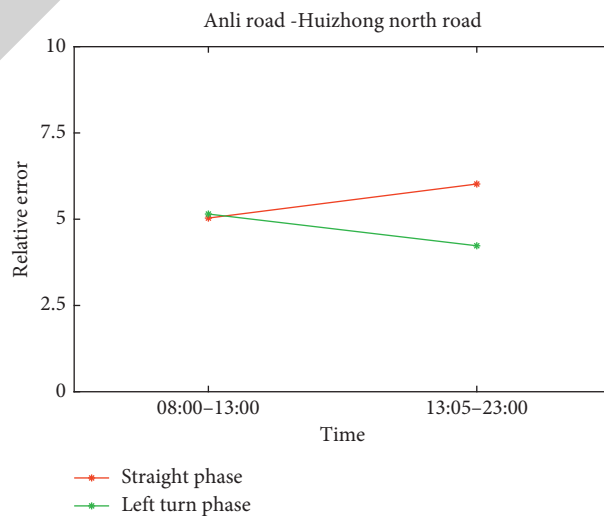
(c)

FIGURE 17: (a) The red time estimation Kehui road-Lincui road. (b) The red time estimation in Anli road-Huizhong road. (c) The red time estimation in Anli road-Huizhong North road.

Huizhong road and Anli-Huizhong North road are slightly larger than Lincui-Kehui road. Therefore, the estimation of results is influenced by the time period within one day.

5. Conclusion

In this paper, the signal cycle length estimation method and the effective red time of intersection estimation are presented based on arriving times of stop line. The time interval of the signal cycle is defined as time_difference; it is the difference between passing stop line times of two vehicles which belong to two different signal cycles. These time differences are distributed periodically that several peaks appear cycle by cycle, and the differences between two neighboring peaks are one cycle length. The method is suitable for calculating three intersections. The results show that the method is achieved successfully. The relative error is around 1% in Lincui road and Kehui road. Most of the estimations are the same as the real cycle length. The red-light estimation algorithm also performed well. The relative error is within the acceptable range. In general, the relationship between sampling rate and error is analyzed based on the above algorithm. The effective peak distribution is verified. The results show that the estimation effects are also related to the number of time intervals each day. The method works well under the following conditions based on our current research:

(1) The number of vehicles passing through intersections in each cycle n/c should be around 0.55. In this case, the number of the time_difference distribution peaks is about 4 after the normal distribution test. The prediction result is more accurate.

(2) The time range of cycle length estimation and red time should not be too long. Compared with the results of two intersections, the result of four time periods is better than that of two time periods.

In future research, the different n/c, time_difference, and other characteristics are considered to improve the two methods; we can also apply related machine learning to our method. These helpful results are applied to various fields of traffic service and automatic driving, etc.

Data Availability

The map data can be found at https://www.openstreetmap.org/#map=4/36.96/104.17. As the navigation data involves data privacy, the data can be acquired by contacting the corresponding author through email.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] X. M. Liu and Z. H. He, "The present situation and trend of development of the urban intelligent traffic system technology," *Automation Panorama*, vol. 2015, no. 1, pp. 58–60, 2015.

[2] P. Hao, X. Ban, K. P. Bennett, Q. Ji, and Z. Sun, "Signal timing estimation using sample intersection travel times," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 792–804, 2012.

[3] Z. Sun and X. Ban, "Vehicle trajectory reconstruction for signalized intersections using mobile traffic sensors," *Transportation Research Part C: Emerging Technologies*, vol. 36, no. 11, pp. 268–283, 2013.

[4] Y. F. Sha and Q. X. Shi, "Research on traffic information collection technology in intelligent transportation system," in *Proceedings of the Academic Seminar on Urban Traffic on Both Sides of the Taiwan Straits*, pp. 331–333, Changhua, Taiwan, 2005.

[5] J. Y. Tan, K. L. Yin, M. Li et al., "Signal timing estimation using mobile navigation data," *Journal of Transportation Systems Engineering and Information Technology*, vol. 17, no. 2, pp. 61–67, 2017.

[6] J. T. Ren, Q. Q. Xie, and J. Yin, "Traffic flow time series separation methods," *Computer Applications*, vol. 25, no. 4, pp. 937–939, 2005.

[7] X. Jeff Ban, H. Ryan, H. Peng et al., "Delay pattern estimation for signalized intersections using sampled travel times," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2130, no. 1, pp. 109–119, 2009.

[8] K. Kwong, R. Kavaler, R. Rajagopal, and P. Varaiya, "Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 6, pp. 586–606, 2009.

[9] M. Kerper, C. Wewetzer, and M. Mauve, "Analyzing vehicle traces to find and exploit correlated traffic lights for efficient driving," *Education & Culture*, vol. 28, no. 3, pp. 310–315, 2012.

[10] S. A. Fayazi, A. Vahidi, G. Mahler, and A. Winckler, "Traffic signal phase and timing estimation from low-frequency transit bus data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 19–28, 2015.

[11] H. Zhao, Q. Chen, W. Shi et al., "Stability analysis of an improved car-following model accounting for the driver's characteristics and automation," *Physica A: Statal Mechanics and its Applications*, vol. 526, pp. 1–12, 2019.

[12] H. Zhao, H. Yue, T. Gu, and W. Li, "CPS-based reliability enhancement mechanism for vehicular emergency warning system," *International Journal of Intelligent Transportation Systems Research*, vol. 17, no. 3, pp. 232–241, 2019.

[13] Y. Tong, L. Zhao, L. Li, and Y. Zhang, "Stochastic programming model for oversaturated intersection signal timing," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 474–486, 2015.

[14] L. Li, X. Su, Y. Wang, Y. Lin, Z. Li, and Y. Li, "Robust causal dependence mining in big data network and its application to traffic flow predictions," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 292–307, 2015.

[15] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, "Trend modeling for traffic time series analysis: an integrated study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3430–3439, 2015.

[16] S. Axer, F. Pascucci, and B. Friedrich, "Estimation of traffic signal timing data and total delay for urban intersections based on low frequency floating car data," in *Proceedings of the Mobil.Tum*, Munich, Germany, July 2015.

[17] S. Axer and B. Friedrich, "A methodology for signal timing estimation based on low frequency floating car data: analysis of needed sample sizes and influencing factors," *Transportation Research Procedia*, vol. 15, pp. 220–232, 2016.

[18] S. A. Fayazi and A. Vahidi, "Crowdsourcing phase and timing of pre-timed traffic signals in the presence of queues: algorithms and back-end system architecture," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 870–881, 2016.

[19] J. Wang, C. Wang, X. Song, and V. Raghavan, "Automatic intersection and traffic rule detection by mining motor-vehicle GPS trajectories," *Computers, Environment and Urban Systems*, vol. 64, pp. 19–29, 2017.