Research on traffic signal cycle optimization based on Webster algorithm

Yuyan Liang [#, *], Qiqi Zhang [#], Zhongchao Wang [#]

Computer Engineering Technical College (Artificial Imtelligence College), Guangdong Polytechnic of Science and Technology, Zhuhai, China, 519090

* Corresponding Author Email: l1952728787@163.com

[#]These authors contributed equally.

Abstract. This study by combining data cleaning, clustering analysis and signal timing method, Webster to famous scenic town traffic control provides a scientific and efficient solution. By removing the missing license plate numbers and abnormal data in the open source data set, the processed data were visualized and analyzed. The cluster analysis was used to divide the peak period of traffic flow, the statistics of the traffic flow period data, the Webster method was used to optimize the timing of traffic lights, and the service level model was established to analyze the results. It was found that it could effectively improve the traffic efficiency. Ease traffic congestion. This research is not only of great significance in theory, but also has application value in actual traffic management, showing its innovation and practical value. The future research of algorithms and to further explore the more advanced technology, to achieve more accurate and efficient traffic flow control.

Keywords: Cluster Analysis, Webster Method, Signal Timing.

1. Introduction

Since the new era, with the rapid development of economy and science and technology, the pace of urbanization continues to accelerate, the popularity of motor vehicles continues to increase, resulting in the deepening of urban road traffic congestion. Especially in the peak tourist season and similar "May Day" and other holidays, some towns with well-known scenic spots, the traffic pressure around the scenic spot is more prominent, providing scientific and reasonable solutions to solve the traffic congestion around the scenic spot is extremely important [1]. Therefore, it is necessary to explore and implement scientific and reasonable traffic management strategies to effectively alleviate or even solve the traffic congestion problem around scenic spots, which is not only related to the image of the city and the quality of life of the residents, but also the key to promote the sustainable development of the tourism industry and promote social harmony and stability.

The Webster method originated in the mid-20th century. With the acceleration of urbanization, the problem of traffic congestion has become increasingly serious. Traffic engineers and researchers needed more effective tools to manage increasingly complex traffic flows, especially at intersections and for signal control. This method proposed by Norman Webster aims to optimize traffic signal cycle and phase settings through mathematical models, which in turn improves traffic efficiency. Wang Shuang [2] et al. used the center street of Danbei Town as the engineering background, and applied Webster's principle to propose an improvement scheme of single-point timing and coordinated control of arterials for intersections, which resulted in an optimal improvement scheme for the overall operation of the road. Bo Wu [3] and others use Webster algorithm to optimize the queuing delay more obviously.

Based on the good foundation of previous studies, this study aims to further follow and extend their excellent results. Specifically, this study will focus on the traffic congestion problem around the scenic area, through the data statistics and the in-depth application of the Webster method, to derive the traffic flow of the intersection of Jingzhong Road - Weizhong Road in each time period; and to optimize the timing of the intersection, and to derive the ratio of the red and green cycles in the case of a reasonable queuing delay time. The purpose of this study is not only to further improve the traffic efficiency and reduce the queuing delay on the basis of previous studies, but also to combine the special traffic demand around the scenic spots, to propose more targeted and practical solutions, and to provide useful references for the traffic management in similar areas.

2. Modeling of the solution

2.1. Data sources

Data sources: https://www.mcm.edu.cn/. As shown in table 1 below:

Table 1. Partial presentation of data sources.

Direction	Time	license plate number	Intersection
3	2024-04-03T14:39:08.632	AF5B7CEM	Huanxi Road - Weizhong Road
1	2024-04-03T17:45:32.316	BK2IA84	Huanxi Road - Weizhong Road
3	2024-04-01T11:47:49.391	CBA7KCG	Huanxi Road - Weizhong Road
2	2024-04-03T18:19:15.892	AFB9CO6	Huanxi Road - Weizhong Road
1	2024-04-01T20:30:49.650	AF25DA6M	Huanxi Road - Weizhong Road
4	2024-04-01T15:43:28.564	AF8CB6CM	Huanxi Road - Weizhong Road

There are 8,844,996 traffic flow data in the data source. It contains columns: direction number, time, license plate number, and intersection where it is located, where direction numbers 1, 2, 3, and 4 correspond to the directions: east to west, west to east, south to north, and north to south. Hereafter collectively referred to as: east inlet, west inlet, south inlet, north inlet.

2.1.1. Data pre-processing

First of all, the preprocessing of traffic flow data, screening found that a total of more than 210,000 data vehicle number is “no license plate”, missing license plate data, as shown in Table 2 below:

At the same time, there are more than 10,000 data license plate number: “unnnowk”, this data is recorded every hour of the day, and this study judges it as abnormal license plate number data, as shown in Table 3 below:

Table 2. Missing data on license plate numbers.

Number of records without license plates: 215723				
No.	Direction	Time	license plate number	Intersection
6008117	3	2024-04-03T15:43:09.229	No license plate	Huanxi Road - Weizhong Road
6008118	2	2024-04-01T12:11:05.237	No license plate	Huanxi Road - Weizhong Road
…	…	…	…	…
8844993	2	2024-05-06T07:30:43.000	No license plate	Jingzhong Road - Huannan Road
8844994	2	2024-05-06T14:12:02.000	No license plate	Jingzhong Road - Huannan Road

Table 3. License plate number anomaly data.

Number of abnormal license plate records: 16872				
No.	Direction	Time	license plate number	Intersection
811871	4	2024-04-07T03:40:16.000	unnnowk	Jingzhong Road - Weizhong Road
2049930	4	2024-04-24T09:16:26.000	unnnowk	Jingzhong Road - Weizhong Road
…	…	…	…	…
8806980	1	2024-05-06T03:57:01.000	unnnowk	Jingzhong Road - Huannan Road
8807325	2	2024-05-06T06:31:10.000	unnnowk	Jingzhong Road - Huannan Road

Due to the large amount of raw data, totaling more than 8 million pieces of data, missing values and outliers were directly eliminated in this study. Since the amount of data after data preprocessing is still relatively large and complex, this study analyzes the data first. Combined with the relevant data, considering that the time period from May 1 to 6 is the Golden Week, most people will go out to travel, resulting in more traffic than usual, so this study only considers the traffic data from April 1st to 30th.

2.1.2. K-Means clustering

K-Means clustering [4] is an algorithm that divides the data points into K clusters such that the variance within clusters is as small as possible and the variance between clusters is as large as possible.

In this study, time t is chosen as the K-value feature to determine the division of the day into time periods, and vehicle counts x_{ij} in different directions are used as the flow characteristics. K data points are randomly selected as the initial centroids, and the distance of each data point from each centroid is calculated, while it is assigned to the cluster represented by the nearest centroid, and then the centroid of each cluster is continuously updated and calculated until the centroid is no longer changing or a predetermined number of iterations is reached. The specific calculation formula is as follows:

Distance calculation:

$$d(x,c) = \sqrt{\sum_{j=1}^{K}(x_{ij} - c_j)^2} \tag{1}$$

Updating center point:

$$c_j = \frac{1}{|c_k|}\sum_{x_i \in c_k} x_{ij} \tag{2}$$

Traffic flow at various times of the day:

$$Flow(c_k) = \sum_{x_i \in c_k}\sum_j x_{ij} \tag{3}$$

Where, c_j denotes the center point of the cluster, d denotes the Euclidean distance, x_{ij} denotes the number of vehicles, $|c_k|$ denotes the number of data points belonging to cluster c_k, $\sum_{x_i \in c_k} x_{ij}$which is the summation of the values of all the data points in the early jth dimension in the cluster c_k, and $Flow(c_k)$ denotes the volume of traffic at each time.

2.1.3. Hierarchical Clustering

Hierarchical clustering [5] is a clustering method that combines or splits data step-by-step and is suitable for situations where the number of clusters is uncertain. The detailed steps can be summarized as follows:

Nitialization: this study initializes the data by considering each data point as a cluster and then finding and merging the two clusters with the closest distance. The calculation of the distance between the new clusters is repeated continuously until the loop stops when all data points are merged into a single cluster or when a predetermined number of clusters is reached.

Step 1 Distance Matrix: Calculate the distance between all the data points, here Euclidean distance is used for calculation.

$$D(i,j) = \sqrt{\sum_{k=1}^{p}\left(x_{ik} - x_{jk}\right)^2} \tag{4}$$

Step 2 Calculate the sum of squares of clusters: For calculating the sum of squares of each cluster, where is the center of mass of the cluster.

$$S(C_i) = \sum_{x \in C_i} \parallel x - \bar{x}_i \parallel^2 \tag{5}$$

Step 3 Merge clusters: Select two clusters and minimize the increment of the sum of squares after merging.

$$\Delta S = S\left(C_{ij}\right) - \left(S\left(C_i\right) + S\left(C_j\right)\right) \tag{6}$$

The sum of squares of the new clusters after merging is:

$$S\left(C_{ij}\right) = S\left(C_i\right) + S\left(C_j\right) + \frac{n_i n_j}{n_i+n_j} D(i,j) \tag{7}$$

Step 4 Update distance matrix: update the distance between the new cluster and other clusters after merging.

$$D\left(C_{ij}, C_k\right) = \sqrt{\frac{n_{ij} n_k}{n_{ij}+n_k} D(i,k) + \frac{n_i n_j}{n_i+n_j} D(i,j)} \tag{8}$$

After that, steps 4 and 5 are repeated until all data points are combined into one cluster or the desired number of clusters is reached.

According to the above formula, this study uses K-means clustering algorithm and hierarchical clustering algorithm to cluster the traffic flow in a day, and then observes the clustering results of different levels to classify the time periods with similar traffic flow into the same category, and finally visualizes the results as shown in Figure 1 below:

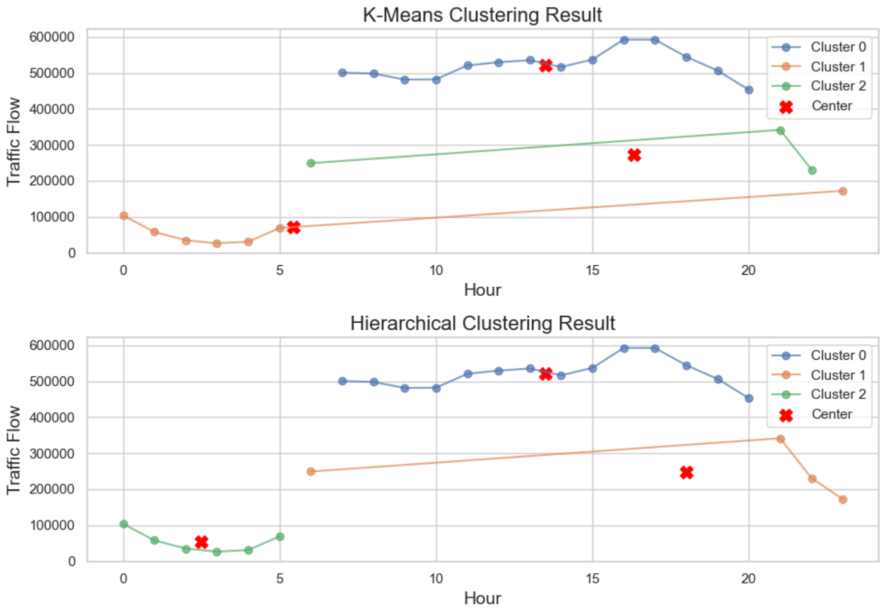

Figure 1. K-means clustering and hierarchical clustering results.

As can be seen from Figure 1, the results of K-means clustering and hierarchical clustering are approximately the same. Therefore, in this study, 7 to 10 o'clock is classified as the morning peak, 11 to 15 o'clock is classified as the afternoon peak, 16 to 20 o'clock is the evening peak, and the rest of the time is the off-peak time. Also to verify the effectiveness of the clustering results, this study evaluated the results of K-means clustering and hierarchical clustering, which are shown in Table 4 below.

Table 4. Assessment indicators for clustering results.

Assessment metrics	Assessment metrics for K-means clustering	Assessment metrics for hierarchical clustering
The Contour Coefficient:	0.734	0.761
The Calinski-Harabas Index:	234.688	260.229
The Davies-Bouldin Index:	0.374	0.334
The Dunn Index:	0.406	0.402

In this study, the clustering effectiveness of K-means clustering and hierarchical clustering at traffic flow time points was evaluated by analyzing the contour coefficient [6], Calinski-Harabasz index [7], Davies-Bouldin index and Dunn index [8]. The results show that hierarchical clustering outperforms K-means in most of the indices, indicating better clustering. Although K-means is faster in processing large-scale data, it is sensitive to the initial center, while hierarchical clustering provides richer structural information and interpretability. Overall, this study confirms the superiority of the clustering effect.

Finally, this study visualizes the obtained data for morning peak, afternoon peak and evening peak and the results are shown in Figure 2 below:

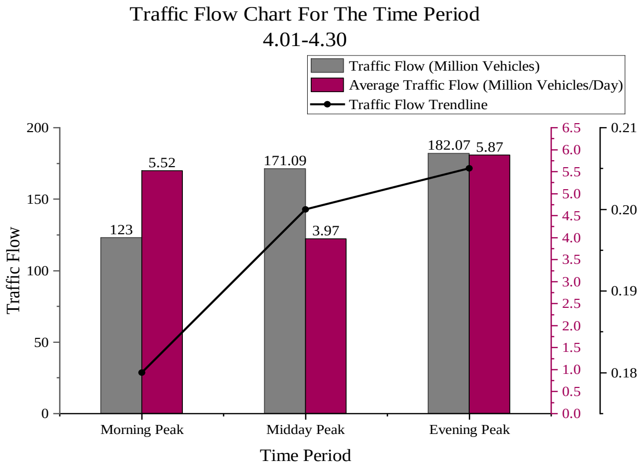

Figure 2. Peak Hour Traffic Flow Map.

As can be seen from Figure 2 the afternoon peak and evening peak hours around the scenic area of the road traffic is larger, the afternoon peak hour traffic vehicles may be due to the midday hours of tourists out of the scenic area to go to the surrounding restaurants caused by the evening peak hour traffic may be due to the scenic area around the residents commute and scenic area tourists to go home caused by the large.

2.1.4. Data screening and results analysis

To sum up, this study visualizes the traffic flow data from April 1 to 30 on a weekly basis, and finds that the traffic flow data on Monday and Tuesday are in the top two. Then the traffic flow data is divided into morning peak, afternoon peak, evening peak and off-peak hours by clustering. Then the traffic flow data of different periods were visualized. For this study, Jingzhong Road-Weizhong Road intersection has the advantage of geographical location, which is in the middle of each intersection, and can reflect the trend of the traffic flow of the entire traffic hub more comprehensively and objectively, and reduce the impact of considering the global intersection contingency factors. Therefore, the traffic flow data of the intersection between Zhonglu and Weizhong Road are screened from the traffic flow data of different segments. The results are shown in Figure 3 below:

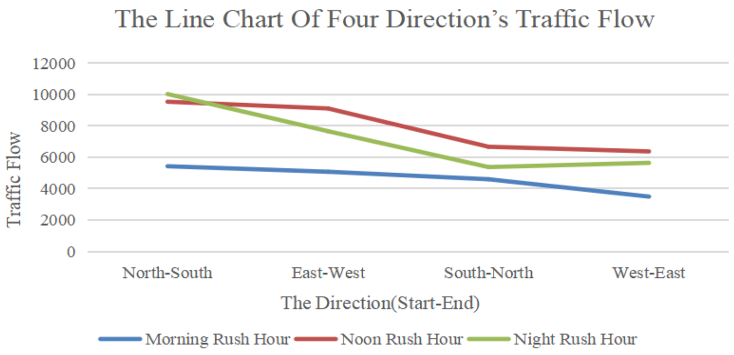

Figure 3. Line graphs corresponding to time periods in the four directions.

Since this study found missing data for April 1 through April 18 with the direction of the west inlet, this study only considered traffic flow data from April 18 through 30, and then counted the average hourly traffic flow for each phase based on the traffic flow data for the Jingzhong Road-Weizhong Road intersection filtered by the time period. This is shown in Table 5 below:

Table 5. Table of average traffic flow for each time period.

Inlet Road	Morning Peak	midday peak	Evening Peak	Off-Peak
East Entrance	463	597	514	254
West Entrance	297	384	368	169
South Entrance	430	436	359	165
North Entrance	500	630	656	282

3. Modeling and Solution

3.1. Webster's Algorithm

3.1.1. Webster timing method

(1) Basic saturation flow rate

Saturation Flow Rate is the maximum number of vehicles passing a lane per unit of time under signal control, measured in pcu (equivalent standard vehicles)/h. Various types of import lanes its dedicated phase basic saturation flow S_{bi}, straight lane for 1400~2000pcu/h, take the average value of 1650pcu/h; left-turn lane for 1300~1800pcu/h, take the average value of 1550pcu/h; right-turn lane for 1550pcu/h. Import lanes width of 3.0m~3.5m.

(2) Generalized Correction Factors for Various Types of Lanes

Corrections to the base saturation flow rate are made by considering the effect of traffic flow (e.g. lane width, nature of traffic flow, etc.). Common correction factors include: lane width: wider lanes of traffic is higher; gradient: uphill and downhill on the flow of the impact; traffic type: the proportion of large and small cars (large cars need to reduce the proportion of saturated flow of high); the impact of non-motorized traffic: sidewalks and non-motorized traffic flow will reduce the saturated flow of traffic flow; traffic flow impact: high traffic flow may lead to congestion, reduce the saturated flow of each lane; correction formula as follows:

$$S = S_{bi} \times f_{\omega} \times f_p \times f_t \times f_b \times f_q \tag{9}$$

Where: S_{bi} is the base saturation flow rate; f_{ω} is the lane width correction factor [9]; f_p is the slope correction factor, which takes the value of 0 in this study. f_t Is the traffic type correction factor; f_b is the non-motorized traffic impact correction factor; and f_q is the traffic volume impact correction factor.

(a) Lane Width Correction:

$$f_{\omega} = \begin{cases} 0.4(W - 0.5) & 2.7 \le W \le 3.0 \\ 1 & 3.0 < W \le 3.5 \\ 0.05(W + 16.5) & W > 3.5 \end{cases} \tag{10}$$

Where: f_{ω} is the lane width correction factor; W is the lane width (m). In this study, the intersection configuration lane type is considered to be three lanes, which are left-turn lane, straight lane and right-turn lane, assuming that each lane $W = 3.5$, so, $f_{\omega} = 1$.

(b) Traffic Type Correction:

$$f_t = \begin{cases} 1.0 & \mathrm{t} > 80 \\ 0.9 & 60 < t \le 80 \\ 0.8 & 40 < t \le 60 \\ 0.7 & t \le 40 \end{cases} \tag{11}$$

Where: f_t is the correction coefficient of traffic type; t is the proportion of small cars. This study is an urban suburban highway, considering this intersection near the town scenic area, there will be more tourists visiting by bus, so assuming that the proportion of small cars is 80% and the proportion of large cars is 20%, then f_t takes the value of 0.9.

(c) Non-motorized traffic impact correction:

$$f_b = \begin{cases} 1.0 & \mathrm{b} < 10 \\ 0.9 & 10 \le b < 25 \\ 0.8 & 25 \le b < 50 \\ 0.7 & b \le 50 \end{cases} \tag{12}$$

Where: f_b is the non-motorized traffic impact correction coefficient; b is the proportion of non-motorized traffic. From the geographical location of the intersection, assuming that the proportion of non-motorized traffic is 20%, f_b takes the value of 0.9.

(d) Traffic flow impact correction:

$$f_q = \begin{cases} 1.0 & \mathrm{q} \le 200 \\ 0.9 & 200 < \mathrm{q} \le 400 \\ 0.8 & 400 < q \le 600 \\ 0.7 & q > 600 \end{cases} \tag{13}$$

Where: f_q is the traffic flow impact correction coefficient, the size of the traffic flow will make vehicles interfere with each other, affecting the ability of vehicle traffic? q Is the size of the traffic flow (pcu/h).

3.1.2. Optimal cycle time

In this study, assuming the same roadway configuration for the cross type intersection Longitudinal Middle Road-Latitudinal Middle Road, the saturated flow rate for each inlet roadway is calculated from equation (8) as follows: east inlet: 1,025 pcu/h; west inlet: 1,154 pcu/h; south inlet: 1,025 pcu/h; and north inlet: 1,025 pcu/h.

Calculate the flow ratio for each phase:

$$y_i = \frac{q_i}{S_i \times N_i} \tag{14}$$

Where: y_i is the i phase flow ratio; q_i is the i phase traffic flow (pcu/h); S_i is the i phase saturated flow rate (pcu/h); N_i is the number of lanes of the i phase inlet lane. Let the east inlet straight left turn be the first phase, the west inlet straight left turn be the second phase, the south inlet straight left turn be the third phase, and the north inlet straight left turn be the fourth phase. Combined with the data in Table 4, this study takes the morning peak as an example, and the flow ratio of 4 phases can be calculated from the above equation y_i, $y_1 = 0.151$, $y_2 = 0.086$, $y_3 = 0.140$, $y_4 = 0.163$.

Calculate the total lost time:

$$L = \sum_{i=1}^{n} l_i + AR \tag{15}$$

Where: n is the number of signal phases; l_i is the i phase loss time (s), and AR is the all-red time (s).

The phase loss time consists of two parts.

First, at the beginning of the green light, there is a start-up time loss due to the fact that the vehicles waiting behind the stop line cannot accelerate immediately, which is generally about 2s.

Second, after the end of the green light, there is a short yellow light time, in order to avoid traffic flow conflicts, there is a loss of time to clear the intersection, which is generally 3s.

Therefore, $L = (2 + 3) \times 4 + 0 = 20(\mathrm{s})$

Calculate the optimal cycle length:

$$C_0 = \frac{1.5L+5}{1-\sum_{i=1}^{n} y_i} \tag{16}$$

Where: C_0 is the optimal cycle duration (s); L is the total loss time (s); y_i is the i phase flow ratio. So, from the above equation C_0, we get $C_0 = 76\text{s}$.

Calculate the effective green light duration;

$$G_e = C_0 - L \tag{17}$$

Where: G_e is the effective green light duration (s), C_0 is the optimal cycle duration (s), and L is the total lost time (s).

So, $G_e = 76 - 20 = 56(\text{s})$.

Calculate the effective green light duration for each phase:

$$g_i = G_e \times \frac{y_i}{\sum_{i=1}^{n} y_i} \tag{18}$$

Where: g_i is the effective green light duration (s) for phase i, G_e is the effective green light duration (s), and $y_i$is the flow rate ratio for phase i. So, from the above equationg_i, we get $g_1 = 15\text{s}$, $g_2 = 9\text{s}$ $g_3 = 15\text{s}$ $g_4 = 17s$

Calculate the green display time for each phase:

$$G_i = g_i - A + l_i \tag{19}$$

Where: G_i is the i phase green light display time (s)[10],A for the yellow light time (s), generally take the value of 3s, g_i for the i phase effective green light time (s), l_i for the i phase loss time (s). So, from the above equation G_i, we get $G_1 = 17s$ $G_1 = 11\text{s}$ $G_1 = 17\text{s}$ $G_1 = 19\text{s}$.

The minimum green time for each phase timing setting in a signal cycle must be able to guarantee the safe pedestrian crossing requirements. The calculation model is as follows:

$$t_p = 7 + \frac{L_p}{v_p} - I \tag{20}$$

Where: t_p is the time for pedestrians to pass through the sidewalk, L_p is the length of the sidewalk, and L_p is taken as 11 m, v_p is the step speed of pedestrians passing through the sidewalk, and is generally taken as 1.2 m/s; I is the effective green light interval time, and I is taken as 5 s. Therefore,$\mathrm{t_p} = 7 + \frac{10.5}{1.2} - 5 \approx 11\text{s}$ $\mathrm{G_i}$ all satisfy the minimum green light time.

3.1.3. HCM Percentage Delay Model

Based on the intersection delay model of the HCM specification of the U.S. Department of Transportation, the Percentile Delay Method (PDM) calculation method model is proposed, which in turn calculates the intersection delay. The model is as follows [11]:

$$D = \frac{V_{D_{10}}+V_{D_{30}}+V_{D_{50}}+V_{D_{70}}+V_{D_{90}}}{(v_{10}+v_{30}+v_{50}+v_{70}+v_{90})\times C} \tag{21}$$

Where: V_{D_i} is the vehicle delay of each signal cycle under the i condition (i takes the value of 10%, 30%, 50%, 70%, 90%, respectively); v_i is the adjusted traffic flow under the i condition (i takes the value of 10%, 30%, 50%, 70%, 90%, respectively); and C is the length of the signal cycle at the intersection.

The vehicle delay V_{D_i} for each cycle is calculated as follows:

$$V_{D_i} = \frac{v_i \times (C-G)^2}{2(1-v_i/s)} \tag{22}$$

Where: G is the green time; s is the saturation flow rate.

The percentage-adjusted traffic flow v_i is calculated as follows:

$$v_i = v + \left[z_i \times \sqrt{v \times C/3600}\,\right] \times \frac{3600}{C} \tag{23}$$

Where: z_i is the i condition (i takes the values of 10%, 30%, 50%, 70%, and 90%, respectively), and -1.28, -0.52, 0, 0.52, and 1.28, respectively; and v is the traffic volume.

3.1.4. Level of Service Modeling

Following the definition of level of service in the U.S. Roadway Capacity Manual 2010[12], control delays continue to be used to evaluate intersection level of service as shown in Table 6 below:

Table 6. Intersection Level of Service Table.

Level of service	Control delay/s	Traffic conditions
A	<10	Unclogged
B	10~20	Slightly congested
C	20~35	No major congestion
D	35~55	No congestion under normal conditions
E	55~80	Just right congestion
F	>80	Overloaded, congested

3.1.5. Optimization results and analysis

The timing optimization of the morning peak signal at the intersection Jingzhong Road-Weizhong Road using Webster's method is shown in Table 7 below:

Table 7. Optimized Timing Data Table for the Morning Peak.

Inlet road	Green time/s	Delay time/s	Level of service
East Entrance	17	47.2	D
West Entrance	11	39.6	D
South Entrance	17	44	D
North Entrance	19	48.2	D

As can be seen from Table 7, in Table 4 traffic flow data using the Webster method of traffic light timing, in the case of normal traffic conditions, basically no congestion, if the traffic flow in other factors under the influence of fluctuating changes, still need to re-optimize the timing to play a better traffic control status.

4. Conclusion

This research will be carried out after pretreatment of data through clustering analysis, according to the size of cars that different times of the day into the morning rush, noon peak and late peak and off-peak hours four times. The average hourly traffic flow of each phase at the intersection of Jingzhong Road and Weizhong Road in each time period is calculated. Webster method is used to obtain the timing optimization and the best cycle length of the morning peak traffic lights at intersections between Jingzhong Road and Weizzhong Road. HCM percentage delay model and service level model are established to calculate the optimized service level of intersection delay and timing respectively for evaluation. The results show good service level.

There are idealized assumptions in this study, and the effects of natural factors, holidays and other special circumstances on traffsic control are not considered. Therefore, in practical applications, the relationship of other factors such as actual flow rate and traffic conditions must be considered in order to more accurately assess the performance of roads and plan transportation facilities. The optimal path configuration in more complex traffic control situations can be solved by intelligent optimization algorithms. With the development of artificial intelligence and big data technology, the Webster method is expected to be combined with intelligent transportation systems to promote sustainable development. Realize dynamic signal control and real-time adjustment of signal timing.

References

[1] Yang Huanhong, Chai Lei, Huang Wentao, et al. Collaborative optimization of dual source tram and electric bus considering road traffic information [J]. Power System Automation, 2024, 48 (17): 77-87.

[2] Wang Shuang, Chen Weiwei, Liu Menglong, et al. Optimization of Road Intersection Organization and Signal Control in Small and Medium sized Towns Based on Webster's Principle [J]. Engineering Technology Research, 2022, 7 (19): 7-9.

[3] Bo Wu, Gong Xu, Luo Lvtian, et al. Optimization effect of Webster's algorithm on queue delay under different traffic flows [J]. Transport Manager World, 2024, (01): 164-166.

[4] Ikotun A M, Ezugwu A E, Abualigah L, et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data [J]. Information Sciences, 2023, 622: 178-210.

[5] Sun Jigui, Liu Jie, Zhao Lianyu. Research on Clustering Algorithm [J]. Journal of Software, 2008, (01): 48-61.

[6] Zhu Lianjiang, Ma Bingxian, Zhao Xuequan. Cluster validity analysis based on contour coefficient [J]. Computer Applications, 2010, 30 (S2): 139-141+198.

[7] Jain a K, Murty M N, Flynn P J. Data clustering: a review [J]. ACM computing surveys (CSUR), 1999, 31(3): 264-323.

[8] Manochandar S, Punniyamoorthy M, Jeyachitra R K. Development of new seed with modified validity measures for k-means clustering [J]. Computers & Industrial Engineering, 2020, 141: 106290.

[9] Ba Tianxing, Meng Ziyue, Yu Shize. Application research on intersection timing optimization method based on Webster's method [J]. People's Public Transport, 2024, (12): 49-51.

[10] Wu Luxiang, Guan Xingquan, Yan Lei, et al. Research on Intersection Signal Timing Optimization Based on VISSIM Simulation [J]. Logistics Engineering and Management, 2021, 43 (12): 35-38.

[11] Huang Wenman. Signal optimization of typical cross shaped intersections based on Synchro system [J]. Transportation and Communications, 2024, 37 (S1): 101-104+112.

[12] Hu Mingwei, Zhai Suyun, Duan Huabo, et al. Comparative Analysis of Low Emission Zones in Europe [J]. Journal of Shenzhen University (Science and Engineering Edition), 2017, 34 (03): 229-237.