

Evaluation and Optimization Method for Urban Signal Control Performance Based on Automatic Vehicle Identification Data

Han Bai

Shandong Jiaotong University

Senyang Xiong

Shandong Jiaotong University

Liang Zhao

Shandong Jiaotong University

Tianzheng Wei

`weitianzheng0220@163.com`

Shandong Jiaotong University

Jian Lu

Shandong Jiaotong University

Xiuguang Wang

Shandong Jiaotong University

Article

Keywords: signal control performance evaluation, signal control optimization, signal timing parameter extraction, Automatic Vehicle Identification data

Posted Date: September 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10816279/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Evaluation and Optimization Method for Urban Signal Control Performance Based on Automatic Vehicle Identification Data

Han Bai¹, Senyang Xiong¹, Liang Zhao¹, Tianzheng Wei^{1,*}, Jian Lu¹, Xiuguang Wang²

¹ School of Transportation and Logistics Engineering, Shandong Jiaotong University, Jinan 250357, China;

² Shandong Zhengqu Traffic Engineering Co., Ltd, Jinan, China.

* Corresponding author: E-mail: weitianzheng0220@163.com

Abstract: To address the issue of insufficient granularity in signal control performance evaluation caused by the difficulty of directly obtaining signal control log data, this paper aims to construct a closed-loop methodology of "control logic inference → fine-grained evaluation → signal optimization" based on Automatic Vehicle Identification (AVI) data, thereby enhancing the refined management level of urban intersection signal control. An integrated inference method for signal control log data, incorporating DBSCAN clustering, Fast Fourier Transform (FFT), and Support Vector Machine (SVM), is proposed. Based on the inferred timing sequences, extraction models for cycle-level evaluation metrics—including saturation degree, queue length (distinguishing between undersaturated and oversaturated states), and average vehicle delay—are established, and Fuzzy C-Means clustering is employed for performance grading. A signal timing optimization model aimed at minimizing delay is formulated, and the Plant Root Growth Optimization (PRGO) algorithm is introduced for solution. The inference error for signal cycle length is less than 5 seconds, and the error for green start time is within 2 seconds. The root mean square error of the queue length model for the east and west approaches is below 3.5 vehicles, respectively, verifying its effectiveness. The PRGO-based optimization scheme reduces the average vehicle delay at the intersection by an average of 9.34%, with a notable decrease in maximum queue length. The proposed method effectively overcomes the bottleneck of missing signal control timing data and enables cycle-level refined evaluation and optimization based solely on AVI data, thereby providing a feasible technical pathway for transitioning urban signal control from "experience-based control" to "data-driven control."

Key words: signal control performance evaluation; signal control optimization; signal timing parameter extraction; Automatic Vehicle Identification data

1 Introduction

With the rapid growth of urban motor vehicle ownership, traffic congestion has become a major constraint on urban development. Among various mitigation strategies, optimizing traffic signal control stands out as one of the most cost-effective and essential measures for enhancing roadway network efficiency. Scientific evaluation of signal control performance serves as the foundation for optimization decisions and relies on two categories of data: traffic flow parameters that reflect vehicle operation states (e.g., volume, speed) and signal control log data that define the control logic itself (i.e., precise timestamps for phases, cycle length, and green start/end times).

Currently, with the large-scale deployment of detection devices such as electronic police

systems, traffic big data represented by Automatic Vehicle Identification (AVI) data have become increasingly abundant, enabling the acquisition of fine-grained traffic flow parameters. However, in stark contrast to the "transparency" of vehicle data, signal control log data—which serve as the benchmark for evaluation—often remain in a "black-box" state. These data are typically stored in closed, heterogeneous signal controller systems and are difficult to access directly due to proprietary communication protocols. This data asymmetry creates a core bottleneck: even with massive amounts of vehicle trajectory data, it remains challenging to conduct genuinely cycle-level and lane-level refined evaluation and optimization without precise alignment with signal control timing sequences. Therefore, inferring the internally invisible signal control logic from widely available external vehicle detection data (AVI) is a critical prerequisite for achieving data-driven refined traffic management.

Effective evaluation and optimization of intersection signal control performance cannot be achieved without the support of signal control log data. In scenarios where timing information is difficult to obtain directly, accurate estimation of signal timing log data becomes a foundational research task. Early methodologies were predominantly based on statistical analysis or probabilistic approaches. HAO^[1] combined traffic flow theory with delay models to estimate timing plans; Ban^[2] estimated signal cycle length using travel time data. To overcome the dependence on data quality, machine learning methods have been introduced. For instance, Genser^[3] utilized Random Survival Forest (RSF) and Long Short-Term Memory (LSTM) neural networks to predict phase durations.

With the increasing prevalence of trajectory data, research on signal timing inference based on Floating Car Data (FCD) has grown substantially. Kerper^[4] was the first to investigate the feasibility of inferring fixed-time signal plans at intersections using FCD; Fayazi^[5] inferred signal timing by observing bus trajectories across consecutive cycles; Protschky^[6] proposed an algorithm that projects trajectories onto a cycle-length range; Yang^[7] developed a cycle length estimation method based on massive GPS trajectory data; An^[8] proposed an iterative optimization model for phase timing parameters using FCD trajectories; Axer^[9] et al. accomplished a trajectory-based signal timing inference method for multiple intersections; Li^[10], based on traffic wave theory, utilized trajectory data to reconstruct the queuing process for cycle length estimation; and Tian^[11] employed grid mapping and clustering algorithms to identify time-of-day periods and cycle lengths.

In recent years, full-sample, fine-grained License Plate Recognition (LPR) data have attracted increasing research attention. Tong Shuai^[12] estimated cycle length and green time using a dual time-domain diagram method; Zhan^[13] proposed a tailored weighted soft-margin maximization model based on LPR data to infer phase parameters; Lin^[14] achieved inference of signal control plans for a single intersection through multi-algorithm fusion; and Yao^[15] formulated the SPaT (Signal Phase and Timing) estimation problem as a bi-level programming model to match phase boundaries with vehicle passage time series. Overall, existing studies either rely on sampled data or focus exclusively on isolated parameters. A complete and fine-grained inversion of the entire intersection timing plan—particularly the inter-phase relationships—remains insufficiently addressed.

Queue length at intersections is a critical control parameter, and its extraction methods can be broadly classified into two categories. The first category involves maximum queue length estimation based on traffic wave theory. Yao^[16] et al. summarized calculation methods for various

complex queue formation patterns; Zhan^[17] et al. proposed a lane-based Gaussian process approximation model for maximum queue length using license plate data; and Wu Hao^[18] established an estimation model by capturing characteristic points of queue formation and dissipation waves. While theoretically rigorous, such methods rely heavily on the assumption of traffic flow homogeneity and exhibit limited adaptability under dynamic and heterogeneous traffic conditions.

The second category involves queue estimation based on the arrival–departure relationship, where the core principle is to infer queuing dynamics through cumulative flow relationships. The classical cumulative curve method proposed by Webster^[19] has been continuously extended. Sharma^[20] utilized detector data to construct arrival–departure curves for distinguishing secondary queues; Zhan^[21] et al. reconstructed lane-based arrival curves using electronic police data to estimate real-time queue length; Feng Yiwen^[22] developed a queue length algorithm based on spatial distribution characteristics derived from Floating Car Data; Wen Xihua^[23] et al. established a data-driven queue estimation model; Luo^[24] et al. determined the number of stops by comparing actual and free-flow travel times to predict maximum queue length; Ju Jianmin^[25] inferred the number of stops by analyzing travel time distribution characteristics and subsequently derived a maximum queue length formulation; Le Huan^[26] employed the Expectation–Maximization (EM) algorithm to fit a Gaussian Mixture Model and estimated the number of stops based on travel time distribution; Zhang Weibin^[27] et al. estimated queue length distribution using historical trajectory data from connected vehicles; and Wei^[28] et al. adopted an empirical Bayes method to estimate arrival rates and solve for queue parameters. Although this category of methods aligns more closely with practical conditions, it often falls short in capturing the detailed spatial evolution dynamics of queues.

In terms of solution algorithms, metaheuristic methods such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) have been widely adopted. Wang Qingsong^[29] et al. employed improved NSGA-II and PSO algorithms, respectively, to solve multi-objective models; Qi Xianfeng^[30] et al. applied the PSO algorithm to address multi-objective formulations. To enhance performance, Xu^[31] et al. integrated PCA-LSTM prediction with Multi-Objective Particle Swarm Optimization (MMOPSO) for signal optimization; Zhao Yu^[32] et al. utilized a GA with an improved selection operator; Wu Yuechen^[33] et al. pioneered the use of the Dandelion Algorithm for solving multi-objective signal timing models; Zhao^[34] et al. proposed a phase-duration-aware reinforcement learning method; and Qin^[35] et al. combined Backpropagation (BP) neural networks with Support Vector Machines (SVM) and optimized signals based on fuzzy control. Nevertheless, conventional intelligent algorithms may still encounter issues such as premature convergence and parameter sensitivity.

To overcome the aforementioned limitations, this paper aims to establish a comprehensive closed-loop methodology based on AVI data, encompassing "control logic inference → fine-grained evaluation → signal optimization." First, to address the bottleneck of inaccessible signal control log data, a hybrid framework integrating DBSCAN clustering, Fast Fourier Transform (FFT), and Support Vector Machine (SVM) is proposed to achieve integrated inference of phases, cycle lengths, and green start times. Based on the precisely inferred timing sequences, cycle-level extraction models for queue length—distinguishing between undersaturated and oversaturated conditions—and average vehicle delay are constructed, and Fuzzy C-Means clustering is employed to grade control performance. Finally, a signal timing optimization model is formulated

with the objective of minimizing delay, and the Plant Root Growth Optimization (PRGO) algorithm, which simulates the root growth mechanism of plants, is introduced for solution. Through simulation-based comparisons with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), the superiority of the PRGO algorithm in terms of global search capability and convergence stability is validated.

2 Methods

The data employed in this study primarily consist of Automatic Vehicle Identification (AVI) data, road network topology data, and signal timing data (used solely for validation purposes). The AVI data were collected from 18 intersections along roadways such as Lashanhe East Road, Binzhou Road, and Jingshi Road in Jinan City, comprising a total of 4,457,476 records. The core fields include anonymized license plate numbers, timestamps, lane identifiers, and intersection names. The road network topology data define the mapping relationships between lanes and turning movements, which are utilized to normalize the AVI data into lane groups. The signal timing data contain information on phases, green durations, and cycle lengths, and are exclusively employed to verify the accuracy of the inferred results. Following the removal of anomalous and redundant records, the data quality is deemed satisfactory.

2.1 Inference of Signal Control Log Data

Accurate signal control log data, such as precise green start times, constitute the foundation for conducting cycle-level fine-grained performance evaluation. However, such data are typically difficult to obtain directly. In this study, based on the fundamental principle of "stop on red, go on green" at intersections, AVI data matched with lane function information are utilized to analyze the temporal sequence characteristics of vehicles passing the stop line across different lane groups. This analysis enables the inference of the underlying signal control logic, thereby deducing the phases, cycle length, and precise green times. The core steps of this inference process are detailed in the following sections.

(1) Phase and Phase Sequence Inference Based on DBSCAN Clustering

Although the signal control scheme is predetermined, the initially released signal phase within a specific time period cannot be explicitly identified in advance. Moreover, to more comprehensively complete the inference process of signal control log data, it is necessary to determine the signal phases during that period. The steps of the clustering algorithm are as follows.

To infer the phases and phase sequence, the DBSCAN algorithm is applied to cluster the vehicle passage timestamps for each normalized lane group. The parameters are set based on traffic flow characteristics ($\epsilon=8s$, $MinPts = 3$) to identify dense clusters of vehicle discharges during green intervals. By calculating the temporal overlap degree between vehicle clusters from different lane groups (as defined in Equation 1), lane groups with a high overlap degree are identified as belonging to the same phase. Finally, all phase clusters are sorted by their start times to obtain the phase execution sequence.

$$overlap_{ij} = \left(\frac{|I \cap J|}{|I|} + \frac{|I \cap J|}{|J|} \right) / 2 \quad (1)$$

Where $overlap_{ij}$ denotes the overlap degree between phases I and J, I and J represent the

sets of vehicle discharge timestamps belonging to the clusters of movement flows i and j , respectively, and $|I|$ and $|J|$ indicate the number of elements in sets I and J .

Flows belonging to the same phase are merged, and the phases are clustered. The start times of each clustered phase group are then sorted to infer the phase sequence.

(2) Cycle Length Inference Based on FFT

Since the signal cycle length may vary throughout the day and AVI data are discrete in the time domain, directly identifying cycle changes is challenging. However, in the frequency domain, periodic characteristics manifest as distinct spectral peaks. Therefore, this study employs Fast Fourier Transform (FFT) to infer the signal cycle length by analyzing the frequency-domain characteristics of vehicle passage time series. In this study, the reciprocal of the frequency corresponding to the maximum amplitude (peak) represents the estimated cycle length. If only one cycle length exists within the study period, the frequencies of the peaks exhibit integer-multiple relationships. Conversely, if multiple cycle lengths are present, multiple peaks appear without integer-multiple relationships.

The phase with the highest traffic volume is selected as the algorithm input. The specific procedure is as follows: 1) Select the phase with the maximum traffic flow and convert its vehicle passage times into a second-by-second binary sequence (1 for presence of a vehicle, 0 for absence). 2) Apply a sliding time window (e.g., 30-minute window length, 10-minute step size) to extract subsequences. 3) Perform FFT on each subsequence, and take the reciprocal of the frequency corresponding to the dominant peak in the power spectrum as the estimated cycle length for that interval. 4) Compare the estimated cycle lengths of adjacent intervals; if the change exceeds a predefined threshold (e.g., 5 seconds), a cycle transition point is identified, thereby demarcating distinct cycle control periods.

(3) Determination of Green Start Times Based on SVM

To precisely locate the green start time of a phase, this study formulates the problem as a binary classification task. Vehicle passage times from two adjacent phases (e.g., east–west through and east–west left-turn) are taken as samples and assigned corresponding phase labels. A Support Vector Machine (SVM) is employed to identify an optimal separating hyperplane, whose projection onto the time axis represents the switching moment between the two phases.

Validation is conducted using the morning peak period at the intersection of Jingshi Road and Qingnian East Road as a case study. The comparison between the signal cycle lengths inferred by the aforementioned FFT method and the actual signal timing plan is presented in Table 1. The results indicate that the cycle length inference error is within 5 seconds, and the method effectively identifies the cycle transition occurring at 9:30:37, thereby confirming its accuracy.

Table 1 Cycle length comparison

Time Period	Actual Cycle Length (s)	Inferred Cycle Length (s)
9:00:00-9:30:37	220	225
9:30:37-10:00:00	210	213

As shown in the table, the FFT-inferred cycle length differs from the actual cycle by no more than 5 seconds, demonstrating satisfactory inference performance. Moreover, the FFT-based approach effectively identifies the time point of cycle transition.

An SVM model is constructed using the adjacent phases "east–west through" and "east–west left-turn." A comparison between the model-inferred green start time for the east–west through phase and the actual timing plan reveals that the error in green start time is within 2 seconds,

confirming the precision of the proposed method in determining phase switching moments.

Both the discrepancy in green duration for the east–west through phase and the cycle length error are within 2 seconds, and the inference of the cycle transition point is also highly consistent with the actual timing. Consequently, the proposed inference model can effectively reconstruct the signal control log timeline at intersections, laying a solid foundation for the subsequent development of fine-grained performance evaluation metrics based on AVI data.

2.2 Selection of Single-Point Signal Control Performance Metrics

The objectives of signal control strategies (e.g., delay minimization, queue management) determine the selection of evaluation metrics. AVI data support the calculation of various evaluation metrics, which can be broadly categorized into two types: independent metrics (e.g., traffic volume, capacity, volume-to-capacity ratio) and dependent metrics (e.g., queue length, delay). The obtainable single-point evaluation metrics are presented in Table 2.

Table 2 Classification of single-point evaluation metrics

Metric Classification	Intersection Metrics
Independent	Traffic volume, Capacity, V/C ratio
Dependent	Queue length, Delay

To achieve refined evaluation, this study subdivides both the spatial and temporal scales. Spatially, the analysis is conducted at the level of individual lanes (left-turn, through, right-turn) for each approach based on AVI data, enabling the assessment of traffic operation states for different movement flows. Temporally, the morning and evening peak periods—characterized by pronounced traffic features—are selected, and a single signal cycle is adopted as the fundamental evaluation unit for fine-grained control performance analysis.

2.3 Intersection Saturation Degree Extraction

(1) Traffic Volume Extraction and Analysis

When evaluating the control performance of an intersection, the variation in traffic flow serves as an intuitive reference indicator. In general, common vehicle types on roadways include passenger cars, medium-sized buses, and heavy trucks. During the calculation process, all vehicle types are converted to equivalent passenger car units (PCU) using standard conversion factors.

The temporal characteristics of traffic flow refer to the periodicity and trends exhibited by traffic data over time. Conducting temporal correlation analysis on traffic flow data can reveal underlying patterns across different time scales, thereby assisting in the selection of appropriate evaluation metrics for signal control during different time periods at an intersection. However, AVI data cover the entire day and involve an enormous data volume. Moreover, during certain periods of the day, traffic volumes in the network are relatively low and cannot adequately reflect the traffic characteristics of the intersection. Therefore, by statistically analyzing the total intersection volume, the overall trend of intersection flow can be inferred, which allows for shortening the study period and reducing the sample size of AVI data.

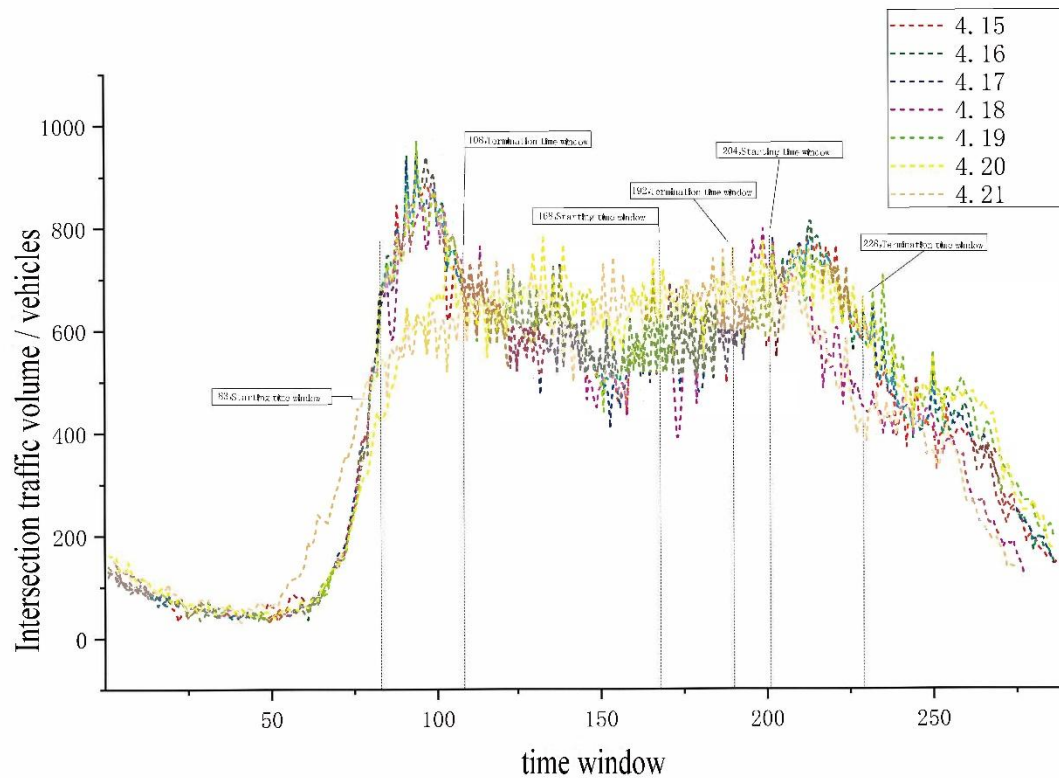
Taking the intersection of Jingshi Road and Qingnian East Road as an example, a day is divided into 288 time windows, and the total number of vehicle passage records within the same time window at the intersection is aggregated to obtain the traffic volume for any given time window. The traffic volumes for selected time windows are shown in Table 3.

Table 3 Hourly traffic volume table

Approach	Time	Volume (veh)
South to North at Jingshi–Qingnian East Rd	2024-04-15 09	1216
North to South at Jingshi–Qingnian East Rd	2024-04-15 15	1101
West to East at Jingshi–Qingnian East Rd	2024-04-15 00	576

As illustrated in Figure 1, the traffic flow on weekdays exhibits a typical bimodal pattern, with the morning and evening peaks aligning closely with residential commuting hours. On the one hand, given that this study aims to evaluate the signal control performance of intersections, it is necessary to select evaluation criteria corresponding to the traffic characteristics of different time periods. The high traffic volumes during weekday morning and evening peaks are more indicative of the intersection's operational characteristics. On the other hand, the pressure on traffic management and control is significantly more pronounced during these peak periods. To ensure efficient data mining and processing for subsequent analysis, this study ultimately confines the temporal scope to the weekday period from 7:30 to 9:30.

Figure 1 Weekly total traffic volume variation at the intersection



(2) Calculation of Saturation Degree

Based on the AVI data, the cycle-level traffic volume for each lane is aggregated, and the lane capacity is calculated in accordance with the Urban Road Engineering Design Specifications. The saturation degree is defined as the ratio of the actual lane volume to the lane capacity. Referring to the U.S. Highway Capacity Manual (HCM), the saturation degree is classified into five levels (A through E), corresponding respectively to free-flow conditions through congested conditions.

2.4 Queue Length Model Development

Queue length is a critical indicator for evaluating signal control performance. Based on the

precisely inferred green start and end times obtained from the previous step, this paper proposes a queue length estimation model that distinguishes between undersaturated and oversaturated conditions.

(1) Undersaturated Condition

When the queued vehicles can be fully discharged during the green interval, the stability of time headways is utilized to identify queued vehicles. For vehicle passage records of a single lane sorted chronologically, the time headway between consecutive vehicles is calculated. The consecutive time headways of queued vehicles typically remain stable within 4 seconds, with the difference between adjacent headways being less than 3 seconds. Once either the headway value or the difference exceeds the respective threshold, the queue is considered to have dissipated, and subsequent arriving vehicles are not counted in the queue. The queue length is then equal to the number of vehicles that pass continuously from the green start time until the moment the queue clears.

(2) Oversaturated Condition

When the queue cannot be fully cleared within a single cycle (indicated by consistently stable time headways throughout the green interval with no abrupt changes), vehicles experiencing secondary queuing must be identified. By matching AVI records of vehicles at upstream and downstream intersections, the travel time of each vehicle is calculated. If a vehicle's travel time during the green interval is significantly greater than the sum of the "red duration of the current cycle + free-flow travel time along the segment," it is identified as having experienced secondary queuing. The queue length is obtained by summing the number of vehicles passing at the saturation flow rate during the green interval and the number of secondarily queued vehicles.

The model is validated using video data captured by an unmanned aerial vehicle (UAV). Taking the intersection of Jingshi Road and Qingnian East Road as an example, the Root Mean Square Error (RMSE) of the queue length for the eastbound through lane (Lane 2) is 3.26 vehicles, and the RMSE for westbound Lane 1 is 1.06 vehicles, indicating satisfactory model accuracy.

2.5 Calculation of Average Vehicle Delay

Intersection delay refers to the difference between the time a vehicle takes to traverse the intersection and its free-flow travel time. The ideal travel time is calculated based on the segment length and the design speed, as shown in Equation (2). The actual travel time is obtained by matching AVI records from upstream and downstream locations. The delay of an individual vehicle is the difference between its actual travel time and the ideal travel time, and the average vehicle delay per cycle is computed according to Equation (3).

The ideal travel time can generally be calculated using the designated travel speed and the distance between two points, as expressed in Equation (2).

$$t_0 = \frac{L}{v_0} \quad (2)$$

where t_0 is the ideal travel time, v_0 is the design speed, and L is the segment length. The average delay is then derived from the difference in travel times, as shown in Equation (3).

$$\bar{t} = \frac{1}{n} \sum_{i=1}^n t_i - t_0 \quad (3)$$

where $\bar{t}$ is the average delay, t_0 is the ideal travel time, t_i is the actual travel time of the i -th vehicle, and n is the total number of vehicles. For the fine-grained, cycle-level average vehicle delay considered in this study, vehicles must be assigned to specific cycle units and individual

lanes to match travel times and compute free-flow times, thereby obtaining the average vehicle delay for each lane under different cycle conditions.

3 Results

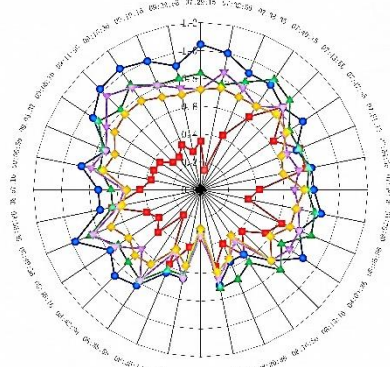
3.1 Single-Point Signal Control Performance Evaluation

The evaluation of signal control performance at signalized intersections typically relies on comprehensive evaluation methods that integrate multiple metrics. These methods convert qualitative indicators into quantitative ones or directly process raw quantitative data to judge the relative performance of the evaluation subject. While numerous comprehensive evaluation approaches exist, they share common limitations: subjectivity in indicator selection and quantitative analysis, potential loss of the original meaning of indicators, and results that are often abstract and difficult to interpret. Moreover, different evaluation methods may yield divergent outcomes due to variations in indicator weighting and assignment schemes. Therefore, this study first conducts a profile analysis to identify specific operational issues within the intersection and subsequently applies Fuzzy C-Means clustering analysis for performance grading.

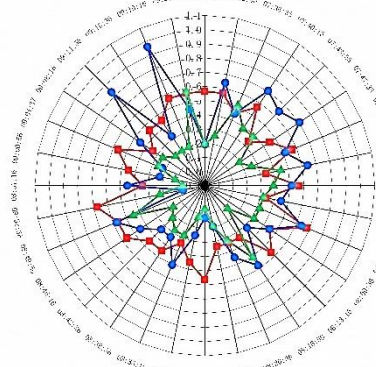
3.1.1 Profile-Based Evaluation and Analysis of Single-Point Signal Control Performance

Based on the extraction methods for saturation degree, queue length, and average vehicle delay presented in Chapter 2, this section analyzes the differences in traffic characteristics—including saturation degree, queue length, and average vehicle delay—across various approach lanes at a signalized intersection, thereby further investigating the underlying signal control issues. According to established classification criteria for signalized intersection evaluation metrics, the performance of different approach lanes within the intersection is assessed. The intersection of Qingnian East Road and Jingshi Road is selected as the study site. Utilizing the inferred signal control log data, traffic parameters at a cycle-level temporal granularity are adopted for evaluation and analysis. A total of 32 cycles within the period from 7:30 to 9:30 are selected for graphical analysis. Since right-turn movements are generally unaffected by signal control, only through and left-turn movements are considered in the analysis.

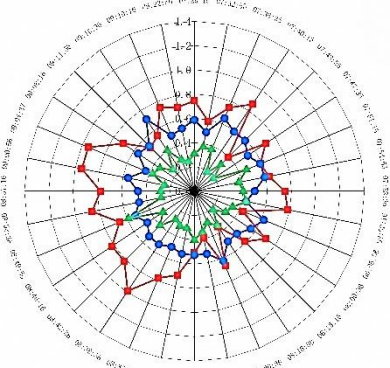
First, the saturation degree is analyzed. The graphical distributions for each approach are presented in Figure 2.



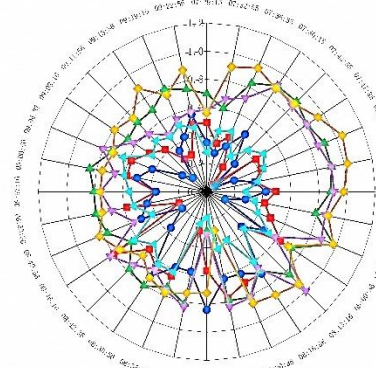
(a) East import



(b) North import



(c) South import



(d) West import

Figure 2 Cycle-level saturation degree distribution for each approach

As can be observed from Figure 2, the saturation degree of through Lane 2 on the east approach is relatively higher compared with other through lanes. After 9:08:16, it remains at Level D, while the saturation degree of the left-turn lane stays at Level B for four cycles and at Level A for the remaining cycles.

On the north approach, both through and left-turn lanes maintain a saturation degree of Level A across most cycles, which is lower than that of other approaches, indicating favorable traffic conditions with minimal congestion.

On the south approach, all through lanes consistently operate at Level A, whereas the left-turn lane exhibits a relatively higher saturation degree. Before 8:38:56, its saturation degree remains at Level C, and for eight consecutive cycles it falls within Levels D and E.

On the west approach, through Lane 4 shows a notably higher saturation degree compared with other through lanes. From the second cycle to the thirteenth cycle, its saturation degree exceeds Level C, and other lanes also exhibit elevated saturation levels during this interval. The left-turn lane, however, remains predominantly at Level A across all cycles.

The cycle-level delays for each lane are further analyzed, as illustrated in Figure 3.

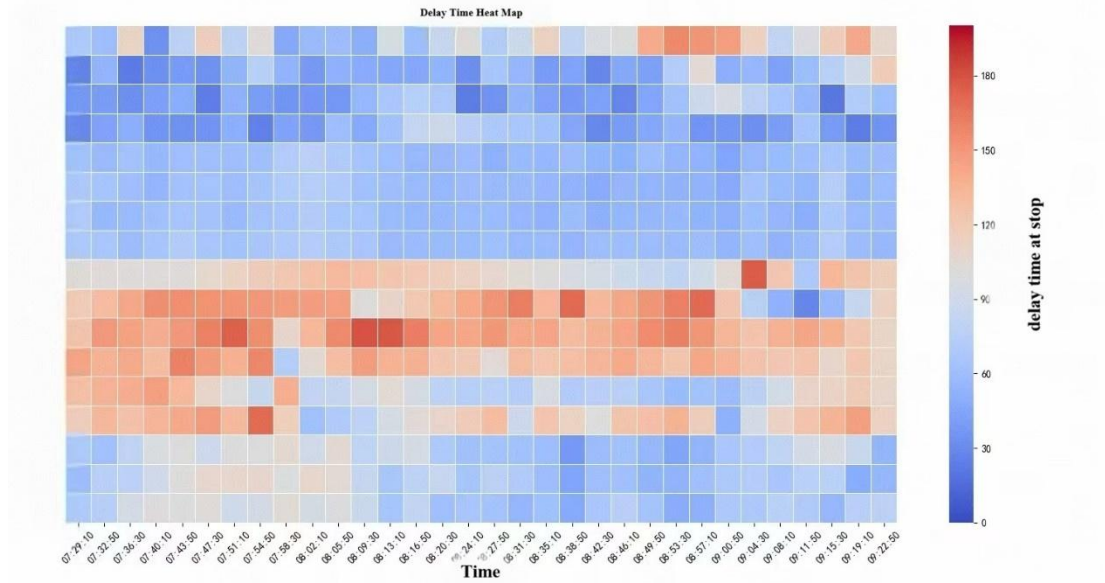


Figure 3 Cycle-level delay distribution for each lane at the intersection

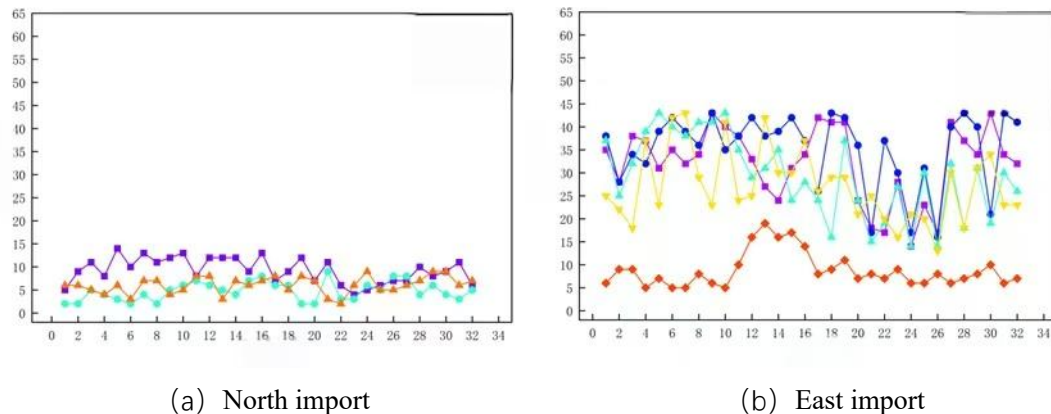
As shown in the figure, the overall delay on the west approach is significantly higher than that on other approaches, with the westbound through lanes exhibiting particularly pronounced delays. The delay values across lanes span a wide temporal range, which may be related to variations in traffic volume over the study period. A few lanes display extreme delay values, reflecting considerable operational fluctuation in that direction.

On the east approach, cycle-level delays are relatively concentrated, indicating stable operating conditions, strong regularity in traffic flow dynamics, and the absence of notable abnormal fluctuations.

The delay patterns observed on the south approach correspond to the characteristics of random fluctuation-induced additional delay described by steady-state theory; namely, overall delay fluctuates within a certain range without sustained excessively high or low values. This suggests that traffic flow on the south approach is relatively stable, albeit subject to random perturbations.

The north approach exhibits low overall delay, with only occasional increases in delay for the northbound left-turn lane during certain cycles. For the majority of cycles, operation remains smooth and traffic pressure is minimal.

The queue length characteristics are further analyzed, as illustrated in Figure 4.



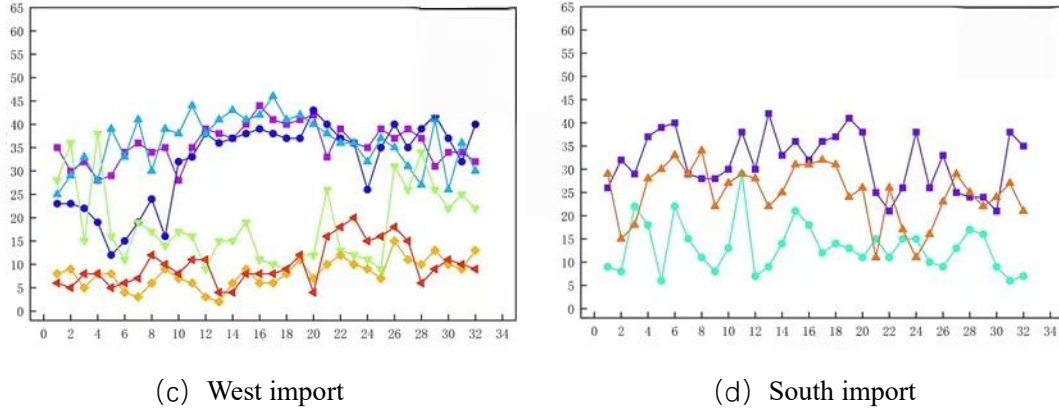


Figure 4 Cycle-level queue length distribution for each lane at the intersection

As shown in Figure 4, the east and west approaches carry the highest traffic volumes, with queue lengths generally exceeding 30 vehicles, indicating both heavy flow and substantial fluctuation. On the east approach, the overall queuing trend is relatively stable from Cycle 1 to 16 but exhibits pronounced volatility from Cycle 17 to 32, with Cycles 20 through 27 showing notably lower queue lengths.

On the west approach, there is a considerable discrepancy in queue lengths between the other through lanes and the westbound through Lane 5, reflecting significant variation in lane utilization. The overall queue lengths on the west approach remain relatively stable, which is also consistent with the higher delay times observed.

The south approach maintains elevated queue lengths, with both the southbound left-turn lane and southbound through Lane 2 queuing approximately 20 vehicles. Given the limited number of lanes, the per-lane traffic pressure is substantial. Moreover, the queue lengths of the two southbound through lanes differ markedly, indicating uneven lane utilization. In contrast, the queue lengths on the north approach remain consistently low.

In summary, it is necessary to rationally allocate green times for the east–west and southbound traffic movements, and to adjust for the saturation and delay conditions observed on the west approach. Optimization of the intersection signal control is therefore warranted.

3.1.2 Control Performance Evaluation Based on Fuzzy C-Means Clustering

Cluster analysis examines the distances and similarities among sample data to partition them into multiple categories. In hard clustering, a sample either fully belongs to or does not belong to a given cluster. In contrast, fuzzy clustering is more flexible: a sample can belong to all clusters, with the sum of its membership degrees across all clusters equaling one. Samples are then classified according to the magnitude of their membership degrees in each cluster. Given that traffic systems are characterized by randomness, nonlinearity, and fuzziness, and that the current definition of traffic control performance entails a certain degree of subjectivity and ambiguity, fuzzy clustering analysis is more suitable for discrimination tasks. Among various fuzzy clustering algorithms, the Fuzzy C-Means (FCM) algorithm is the most widely applied in practice due to its favorable generalization performance.

The Fuzzy C-Means clustering algorithm was first proposed by Dezdek in 1981. The algorithm is driven by an optimization function that seeks to minimize the weighted sum of distances from all samples to the cluster centers, combined with their membership degrees,

thereby classifying samples according to their membership levels. The computational procedure of the algorithm is as follows:

Assume a dataset $X = \{x_1, x_2, \dots, x_n\}$. If the data are to be partitioned into c classes, denoted by $\{A_1, A_2, \dots, A_c\}$, with $2 < c < n$. $V = \{v_1, v_2, \dots, v_c\}$ represent the cluster centers of each class, and let $U = [\mu_{ik}]_{c \times n}$ denote the membership degree matrix, where μ_{ik} indicates the membership degree of the i -th sample with respect to the k -th class.

The objective function can be expressed as Equation (4):

$$J(U, V) = \sum_{i=1}^n \sum_{k=1}^c (u_{ik})^m (d_{ik})^2 \quad (4)$$

where m denotes the fuzzy exponent; d_{ik} is the Euclidean distance from sample x_k to the cluster center v_i , obtained from Equation (5):

$$(d_{ik})^2 = \|x_k - v_i\|^2 \quad (5)$$

subject to the following constraint, as shown in Equation (6):

$$\sum_{k=1}^c \mu_{ik} = 1, \mu_{ik} \in [0, 1] \quad (6)$$

When $J(U, V)$ is minimized—i.e., the distances between samples and their respective cluster centers are minimized—the optimal cluster centers are obtained. The membership degree μ_{ik} and the cluster center matrix v_i are respectively updated using Equations (7) and (8):

$$\mu_{ik} = \left(\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (7)$$

$$v_i = \frac{\sum_{k=1}^c (u_{ik})^m x_k}{\sum_{k=1}^c (u_{ik})^m} \quad (8)$$

Currently, the existing domestic frameworks related to control performance evaluation include the *Urban Traffic Operation Status Evaluation Specification* (GB/T 33171–2016), which clearly defines fundamental operational status indicators such as traffic volume and speed, and classifies the traffic operation conditions of road segments and networks into five levels. Additionally, the *Road Traffic Congestion Evaluation Method* (GA/T 115–2020) issued by the Ministry of Public Security categorizes the traffic states at signalized intersections into different levels. Since this study focuses specifically on the operational effectiveness of intersections under signal control, and most existing standards primarily address the classification and identification of traffic operation states, this paper refers to the aforementioned specifications along with the practical clustering outcomes to classify signal control performance into three distinct levels.

Given the considerable variation in the ranges of data values and the dimensional differences among parameters, which may affect model convergence speed and clustering results, this study applies min–max normalization to preprocess the raw data, as expressed in Equation (9):

$$\hat{x}_i = \frac{(x_i - x_{i,\min})}{(x_{i,\max} - x_{i,\min})} \quad (9)$$

where $x_{i,\min}$ and $x_{i,\max}$ denote the minimum and maximum values of the i -th dimensional data, respectively, and $\hat{x}_i$ represents the normalized value. For the three indicators extracted in this paper—saturation degree, queue length, and average vehicle delay—Fuzzy C-Means clustering is applied to the eastbound through lane and westbound through lane, which were identified in the previous section as exhibiting poor signal control performance. To enhance noise robustness while balancing clustering fuzziness and inter-class separability, the number of clusters c is set to 3, the fuzzy exponent m is set to 1.5, the minimum error tolerance is 1×10^{-11} , and the maximum

number of iterations is 1500. The spatial distributions of the samples and the resulting cluster centers are presented in Figure 5 and Figure 6.

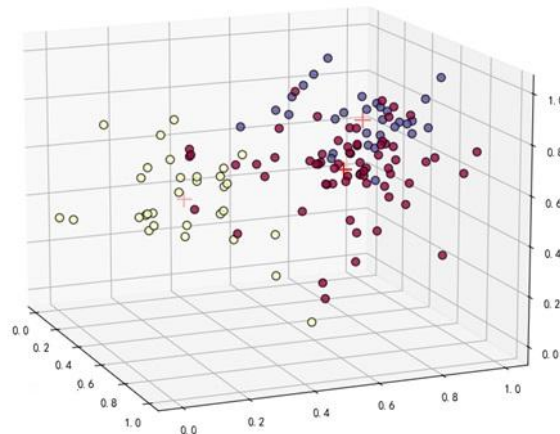


Figure 5 Spatial distribution of samples and cluster centers for the westbound through lane

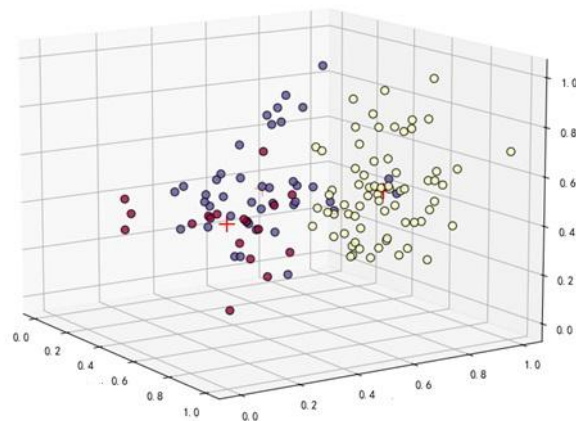


Figure 6 Spatial distribution of samples and cluster centers for the eastbound through lane

The resulting cluster centers are presented in Equations (10) and (11):

$$V_e = [V_1, V_2, V_3] = \begin{bmatrix} 0.49 & 0.74 & 0.79 \\ 52.48 & 53.16 & 56.93 \\ 20.55 & 29.67 & 37.4 \end{bmatrix} \quad (10)$$

$$V_w = [V_1, V_2, V_3] = \begin{bmatrix} 0.47 & 0.71 & 0.87 \\ 91.99 & 130.15 & 145.81 \\ 17.51 & 28.04 & 38.03 \end{bmatrix} \quad (11)$$

where V_e and V_w denote the cluster centers for the eastbound through lane and the westbound through lane, respectively; V_1 , V_2 , and V_3 correspond to the three control performance levels: Level A, Level B, and Level C. The rows of each matrix represent the values of the three indicators—namely, saturation degree, average vehicle delay, and queue length, in that order—

under a given performance level, while the columns represent the indicator values within the same traffic condition. The sample counts corresponding to each control performance level are summarized in Table 4.

Table 4 Distribution of control performance levels

Control Performance Level	Level A	Level B	Level C
The sample size of the eastbound straight lane	37	52	39
Proportion (%)	28.9	40.6	30.5
The sample size of the westbound straight lane	31	37	60
Proportion (%)	24.2	28.9	46.9

Combining the cluster center results with the control performance distribution table, it can be observed that for the eastbound through lane at the intersection, the proportions of cycles classified as Level C and Level A are 28.9% and 30.5%, respectively. This indicates that for approximately one-third of the two-hour period, the control performance is relatively poor, while Level B accounts for 40.6%, representing a predominantly moderate state. For the westbound through lane, the proportions of cycles classified as Level B and Level A are 24.2% and 28.9%, respectively, whereas Level C accounts for 46.9%, occupying nearly half of the time. This demonstrates that the control performance of both the eastbound and westbound through lanes—particularly the westbound through lane—is suboptimal. Through FCM clustering, the control performance state of each cycle within the study period can be clearly identified, thereby facilitating the diagnosis and subsequent optimization of traffic control performance.

3.2 Evaluation of Arterial Coordination Control Performance

The performance of arterial coordination control must be evaluated over a relatively long temporal range. In this study, the green wave segment of Jingshi Road (from Qingnian East Road to Qianfoshan Road) during the morning peak period from 7:30 to 9:30 is selected as the study interval. Based on AVI data, the following metrics are extracted: ① The ratio of the 95th percentile travel time to the average travel time is defined as the Planning Travel Time Index, as expressed in Equation (12):

$$I_p = \frac{TT_{95th}}{TT_{average}} \quad (12)$$

$I_p \approx 1$: The 95th percentile travel time is close to the average value, indicating high operational stability, effective arterial coordination control, minimal fluctuation in vehicle travel times, and no significant extreme delays. $I_p > 1$: The 95th percentile travel time exceeds the average travel time, implying that actual travel times surpass expectations and additional delays are present, indicating suboptimal arterial coordination control. This may result from unreasonable signal timing, traffic conflicts, or other factors that reduce operational efficiency. Higher travel time stability corresponds to better arterial coordination performance, while greater deviation indicates larger travel time fluctuations and a need for optimization of arterial coordination control.

This metric is used to measure travel time stability. ②Traffic Flow Concentration Degree, as expressed in Equation (13):

$$E_i = \frac{\sum_{j=1}^m E_j}{m} = \frac{\sum_{j=1}^m \frac{c_j p_{AOG,j}}{g_j}}{m} \quad (13)$$

where $P_{AOG,j}$ denotes the proportion of vehicles arriving during the green phase at intersection j , C_j is the cycle length, g_j is the green phase duration, m is the number of intersections, and E_i represents the traffic flow concentration degree in direction i . This metric comprehensively reflects the degree of alignment between the green split and the green arrival rate by incorporating the green-to-cycle ratio, thereby providing a reasonable measure of the relationship between cycle length and green phase duration. The calculation results are presented in Tables 5 through 7.

Table 5 Planning Travel Time Index for bidirectional coordination

Direction	Average Travel Time (s)	95th Percentile Travel Time (s)	Planning Travel Time Index
West to East	500.14	850	1.69
East to West	447.44	664	1.48

The green arrival rates are calculated for each intersection along both directions of the arterial coordination segment of Jingshi Road, spanning from Qingnian East Road to Qianfoshan Road. The resulting traffic flow concentration degrees are presented in Table 6.

Table 6 Traffic Flow Concentration Degree for bidirectional coordination

Coordination Direction	Traffic Flow Concentration Degree
East to West	1.06
West to East	0.94

The average speeds of all vehicles traversing the entire segment are presented in Table 7.

Table 7 Average speed for bidirectional coordination directions

Coordination Direction	Average Speed (km/h)	Level
East to West	30.7	Basically Free Flow
West to East	28.7	Mild Congestion

As shown in Table 5, the Planning Travel Time Index for the east-to-west direction (1.48) is lower than that for the west-to-east direction (1.69), indicating better travel time stability in the east-to-west direction. From Table 6, the Traffic Flow Concentration Degree is 1.06 for east-to-west and 0.94 for west-to-east, suggesting a higher degree of alignment between vehicle arrivals and green start times in the east-to-west direction. As presented in Table 7, the average speed in the east-to-west direction is 30.7 km/h (Basically Free Flow), whereas the west-to-east direction averages 28.7 km/h (Mild Congestion). Collectively considering these three metrics, the arterial coordination performance in the east-to-west direction is superior to that in the west-to-east direction.

3.3 Signal Timing Optimization and Simulation Validation

Based on the evaluation results presented in Sections 3.1 and 3.2, a single-intersection signal control optimization model is formulated with the objective of minimizing average vehicle delay. To solve this model, three algorithms—Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and the Plant Root Growth Optimization (PRGO) algorithm—are employed to search for optimal solutions, and their performance and optimization effects are compared through simulation. Additionally, recommendations for optimizing arterial coordination control will be proposed.

3.3.1 Single-Point Signal Control Optimization Model

To optimize signal timing, an objective function is constructed to minimize the total average vehicle delay at the intersection. Among classical delay models, the ARRB model demonstrates favorable accuracy and applicability, and is therefore adopted as the basis for the optimization objective in this study. The objective function is expressed as Equation (14):

$$\min Z = \min[\sum_i^m \bar{D}(C, g)] \quad (14)$$

where $\sum_i^m \bar{D}(C, g)$ represents the average vehicle delay at the intersection calculated using the ARRB model, which is a function of the cycle length C and the green durations g of each phase.

(1) Constraints

1) Effective Green Time Constraint

The effective green time for each phase must comprehensively account for factors such as driver reaction time and pedestrian crossing time to ensure the safe passage of vehicles and pedestrians through the intersection. Typically, the green time should be controlled within a range of 20 seconds to 100 seconds, satisfying the minimum safety requirements while avoiding excessive delays, as expressed in Equation (15):

$$g_{imin} \leq g_i \leq g_{imax} \quad (15)$$

2) Signal Cycle Constraint

To prevent excessive waiting times for vehicles on all approaches, the signal cycle length should be maintained within a reasonable range, typically not exceeding 250 seconds. Meanwhile, to ensure adequate operational efficiency at the intersection, the minimum signal cycle length generally should not be less than 40 seconds, as expressed in Equation (16):

$$C_{min} \leq C \leq C_{max} \quad (16)$$

3) Phase Saturation Degree Constraint

The phase saturation degree is typically related to the operational conditions of the intersection. In this study, the saturation degree is constrained within the range of 0.6 to 0.9, as expressed in Equation (17):

$$\rho_{min} \leq \rho_i \leq \rho_{max} \quad (17)$$

In summary, the signal optimization model for the intersection is formulated as Equations (18) and (19):

$$\min Z = \min[\sum_i^m \bar{D}(C, g)] \quad (18)$$

$$s.t. \begin{cases} g_{imin} \leq g_i \leq g_{imax} \\ \rho_{min} \leq \rho_i \leq \rho_{max} \\ C_{min} \leq C \leq C_{max} \end{cases} \quad (19)$$

(2) Model Solution Based on the PRGO Algorithm

To solve the aforementioned nonlinear optimization model, this study employs the Plant Root Growth Optimization (PRGO) algorithm. The PRGO algorithm simulates the process by which plant roots explore soil for nutrients. By maintaining a balance between search diversity and learning from historically superior solutions, it effectively avoids convergence to local optima, making it well-suited for solving signal timing problems. The steps for applying the PRGO

algorithm to this optimization problem are summarized as follows: 1) Initialize a population of timing plans that satisfy the constraints. 2) Calculate the fitness value (delay) of each plan. 3) According to the PRGO rules, simulate the growth of "fibrous roots" toward regions of optimal solutions and the random exploration of "adventitious roots" to update the population positions. 4) Repeat the iteration process until the stopping criteria are met, and output the optimal timing plan.

3.3.2 Validation of the Single-Point Signal Control Model

(1) Current Signal Control Status at the Intersection

To validate the optimization model, the intersection of Qingnian East Road and Jingshi Road in Jinan City is selected as the study site. Each approach at this intersection is equipped with exclusive turning lanes, and a five-phase signal plan is implemented during peak hours. The specific signal timing parameters are presented in Table 8.

Table 8 Signal timing plan for the Qingnian East Road–Jingshi Road intersection

Signal timing plan for the Qingnian East Road–Jingshi Road intersection (Unit: s)					
Phase No	Movements Included	Green Time	Yellow Time	Cycle Length	
1	East–West Left Turn	29	3	220	
2	East–West Through	93	3		
3	North–South Left Turn	17	3		
4	South Single Release	34	3		
5	North–South Through	32	3		

(2) Experimental Results and Analysis

To solve the optimization model, Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and the Plant Root Growth Optimization (PRGO) algorithm are employed, respectively, to search for optimal solutions. The signal timing plans (phase green durations and cycle lengths) obtained after optimization by each algorithm are presented in Table 9.

Table 9 Signal timing plans for different time periods

Time Period	Scheme	Phase Duration (s)					Cycle Length (s)
		1	2	3	4	5	
7:30-8:30	GA	22	75	28	44	22	206
	PSO	21	80	27	43	22	208
	PRGO	22	79	34	36	22	206
8:30-9:30	GA	22	74	22	36	22	191
	PSO	23	71	22	35	22	187
	PRGO	22	74	23	30	24	188

To verify the effectiveness of the optimized schemes, microscopic simulation is conducted using VISSIM simulation software for both the existing timing plan and the plans optimized by the three algorithms.

Detection locations and evaluation criteria for queue length, delay, and data collection are configured accordingly. Given the inherent differences between real-world traffic flow characteristics and simulated conditions, VISSIM is employed to simulate the existing signal

timing plan as well as the plans optimized by GA, PRGO, and PSO. The schemes are evaluated based on queue length and average vehicle delay, as presented in Table 10.

Table 10 Simulation results

Time Period	Traffic flow direction	Maximum Queue Length (m)				Average Delay (s)			
		Existing	GA	PSO	PRGO	Existing	GA	PSO	PRGO
1	East	194.1	178.3	173.2	169.9	42.8	38	37.5	37.9
	South	162.6	160.6	160.6	152.8	34.1	33.9	36.9	35.8
	west	234.1	224.2	221.5	214.2	25.6	22.6	22.2	20.9
	north	70.1	70	69.5	62.6	64.2	63.7	66.9	64.9
	intersection	234.1	224.2	221.5	214.2	69	65.4	61.9	60.4
2	East	183.9	178.5	169.8	163.4	39.3	36.2	37.7	34.3
	South	164.7	165.5	145.7	157.4	45.7	42.2	42.7	41.5
	west	223.9	228.3	214.1	195.8	30.2	25.5	23.9	25.3
	north	72.7	63.7	59.2	60.8	67.7	67.9	64.5	64.1
	intersection	223.9	228.3	214.1	195.8	57.3	56.5	53.3	52.6

The simulation results indicate that all optimized schemes improve the performance of the intersection. Among them, the scheme optimized by the PRGO algorithm achieves the best performance. Across the two study periods, the average vehicle delay at the intersection is reduced by an average of 9.34%, and the maximum queue length is also reduced to the greatest extent.

4 Discussion

This study successfully achieves accurate inference of signal control parameters and fine-grained performance evaluation based on AVI data, and validates the effectiveness of the PRGO algorithm in signal timing optimization. The following discussion addresses the research findings, methodological limitations, and future directions.

The PRGO-optimized scheme reduces the average vehicle delay on the west approach by approximately 18.4% (from 25.6 s to 20.9 s), a reduction notably greater than that observed on other approaches. This may be attributed to the fact that, under the existing timing plan, the west approach exhibits the highest saturation degree (frequently at Level D or E), thereby constituting the primary bottleneck of the intersection. Through global optimization, the PRGO algorithm allocates additional green time to this direction, yielding higher marginal benefits. In contrast, the north approach inherently operates at a low saturation degree, leaving limited room for improvement and resulting in negligible delay reduction. This finding underscores the necessity of targeted optimization informed by fine-grained evaluation results. Previous studies indicate that single-point signal optimization based on GA or PSO typically reduces delay by 5%–10%. The 9.34% reduction achieved by the PRGO algorithm in this study lies at the upper end of this range, and the 18.4% reduction specifically attained at the west approach bottleneck demonstrates the algorithm's adaptability under asymmetric traffic flow conditions.

Compared with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), PRGO exhibits superior solution quality. We contend that PRGO, by simulating the "gravitropism" and "branching" characteristics of plant root systems, achieves a better balance between global exploration and local exploitation, rendering it less susceptible to premature convergence. Consequently, it is capable of identifying superior solutions within the complex search space of timing plans, thereby offering a competitive new tool for solving similar traffic optimization problems.

The license plate matching rate of AVI data directly influences the accuracy of travel time estimation and the oversaturated queue length model. In the present case study, the data quality is favorable; however, under scenarios with lower matching rates (e.g., nighttime or low-volume periods), model performance may degrade. Future research could explore the integration of multi-source data (e.g., geomagnetic sensors, radar) to enhance model robustness, as well as investigate adaptive adjustment strategies for the parameters of the PRGO algorithm. Furthermore, this study focuses on single-point optimization; subsequent work may extend the proposed methodology to arterial coordination control.

5 Conclusion

To address the challenge of fine-grained evaluation posed by the unavailability of signal control log data, this study proposes a comprehensive methodological framework based on Automatic Vehicle Identification (AVI) data, encompassing "control logic inference → accurate performance evaluation → intelligent signal optimization." The main conclusions are summarized as follows:

A hybrid signal control log data inference method integrating DBSCAN clustering, Fast Fourier Transform (FFT), and Support Vector Machine (SVM) is proposed, enabling the estimation of phases, cycle lengths, and green start times without relying on internal signal controller data. The inference error for cycle length is within 5 seconds, and the error for green start time is within 2 seconds.

Based on the accurately inferred control timing sequences, a cycle- and lane-level fine-grained evaluation metric extraction model is established, with particular emphasis on a queue length estimation model that distinguishes between undersaturated and oversaturated conditions. Case application demonstrates that the proposed method can effectively identify intersection bottlenecks (e.g., the westbound through lane operates at the worst control performance level for 46.9% of the time) and quantify the imbalance in arterial coordination (e.g., the Planning Travel Time Index in the west-to-east direction is 14% higher than that in the east-to-west direction).

A signal timing optimization model with the objective of minimizing delay is formulated, and the Plant Root Growth Optimization (PRGO) algorithm is introduced for solution. Simulation validation indicates that the PRGO algorithm exhibits superior global search capability compared with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The optimized timing plan reduces the average vehicle delay at the intersection by an average of 9.34% and significantly decreases the maximum queue length.

This study confirms the feasibility of leveraging widely available AVI data to drive refined traffic signal control, thereby providing key technical support for the transition from "experience-based control" to "data-driven control."

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Shandong Provincial Natural Science Foundation project (ZR2025QC741), the Inovative Team Project of Ji'nan Government (202333036), Shandong Jiaotong University Talent Introduction Research Start-up Fund (BS2023035).

6. Patents

Author Contributions: Conceptualization, Han Bai and Tianzheng Wei; methodology, Han Bai and Liang Zhao; software, Han Bai and Liang Zhao; validation, Han Bai, Senyang Xiong and Tianzheng Wei; formal analysis, Han Bai and Liang Zhao; investigation, Han Bai and Tianzheng Wei; re-sources, Tianzheng Wei, Jian Lu and Xiuguang Wang; data curation, Han Bai and Liang Zhao; writing—original draft preparation, Han Bai; writing—review and editing, Tianzheng Wei, Jian Lu and Xiuguang Wang; visualization, Han Bai and Liang Zhao; supervision, Tianzheng Wei; project administration, Tianzheng Wei; funding acquisition, Tianzheng Wei. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgments: This work was supported by Shandong Provincial Natural Science Foundation project (ZR2025QC741), the Inovative Team Project of Ji'nan Government (202333036), Shandong Jiaotong University Talent Introduction Research Start-up Fund (BS2023035).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References:

- [1] Ban X, Herring R, Hao P, et al. Delay Pattern Estimation for Signalized Intersections Using Sampled Travel Times [J]. Transportation Research Record, 2009, 2130(1): 109-119.
- [2] Hao P, Ban X, Bennett P K, et al. Signal Timing Estimation Using Sample Intersection Travel Times. [J]. IEEE Trans. Intelligent Transportation Systems, 2012, 13(2): 792-804.
- [3] Genser A, Ambuhl L, Yang K, et al. Time-to-Green predictions: A framework to enhance SPaT messages using machine learning[J]. IEEE, 2020, 10(2): 1-6.
- [4] Kerper M, Wewetzer C, Mauve M. Analyzing vehicle traces to find and exploit correlated traffic lights for efficient driving[J]. Education & Culture, 2012, 28(3): 310-315.
- [5] Fayazi A S, Vahidi A, Mahler G, et al. Traffic Signal Phase and Timing Estimation From Low-Frequency Transit Bus Data. [J]. IEEE Trans. Intelligent Transportation Systems, 2015, 16(1): 19-28.
- [6] Protschky V, Ruhhammer C, Feit S. Learning Traffic Light Parameters with Floating Car Data[J]. IEEE, 2015, 28(3): 2438-2443.

- [7] Yang Q, Yu J, Han J M. Traffic Signals Timing Cycle Length Learning: Using Taxi Gps Trajectories[J]. ICMLC, 2018, 12(3): 13-18.
- [8] An C, Xia J, Lu Z, et al. A new one-way bandwidth-based traffic signal coordination approach based on travel speed variations[C]//Transportation Research Board Annual Meeting. Washington, DC: Transportation Research Board, 2015: 15-3604.
- [9] Axer S, Friedrich B. A Methodology for Signal Timing Estimation Based on Low Frequency Floating Car Data: Analysis of Needed Sample Sizes and Influencing Factors [J]. Transportation Research Procedia, 2016, 15: 220-232.
- [10] LI L, Su X, Zhang Y, et al. Trend modeling for traffic time series analysis: an integrated study[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(6): 3430-3439.
- [11] Xinmei T, Deqi C, Xuedong Y, et al. Estimation Method of Intersection Signal Cycle Based on Empirical Data [J]. Journal of Transportation Engineering, Part A: Systems, 2021, 147(3): 1-12.
- [12] Tong S. Data-Driven Field Timing Plan Estimation and System Development [D]. Zhejiang University of Technology, 2020.
- [13] Zhan X, Li R, Ukkusuri V S. Link-based traffic state estimation and prediction for arterial networks using license-plate recognition data [J]. Transportation Research Part C, 2020, 117.
- [14] Qinghai L, Jinyong C, Guilong L, et al. Signal timing parameters inference method at intersections using license plate recognition data [J]. IET Intelligent Transport Systems, 2022, 16(8): 1092-1107.
- [15] Jiarong Y, Hao W, Keshuang T. A bi-Level programming method for SPaT estimation at fixed-time controlled intersections using license plate recognition data [J]. Transportmetrica B: Transport Dynamics, 2023, 11(1): 1045-1070.
- [16] Yao J, Tang K. Cycle-based queue length estimation considering spillover conditions based on low resolution point detector data[J]. Transportation Research Part C: Emerging Technologies, 2019, 109: 1-18.
- [17] Zhan X, Li R, Ukkusuri S. Link-based traffic state estimation and prediction for arterial networks using license-plate recognition data[J]. Transportation Research Part C: Emerging Technologies, 2020, 117: 102660.
- [18] Wu H, Liu L, Tang K. Queue Length Estimation at Signalized Intersections Based on Ensemble Learning [J]. Journal of Tongji University (Natural Science Edition), 2023, 51(3): 405-415.
- [19] Webster FV. Traffic signal settings[J]. Road Research Technical Paper, 1958, 39. Vermont South, Australia.
- [20] Sharma A, Bullock D M, Bonneson J A. Input-Output and Hybrid Techniques for Real-Time Prediction of Delay and Maximum Queue Length at Signalized Intersections[J]. Transportation Research Record Journal of the Transportation Research Board, 2007, 2035(2035): 69-80.

- [21] Zhan X, LI R, Ukkusuri S V. Lane-based real-time queue length estimation using license plate recognition data[J]. Transportation Research Part C, 2015, 57: 85-102.
- [22] Feng Y. Queue Length Detection at Road Intersections Based on Large-Scale Vehicle Trajectory Data [D]. Shenzhen University, 2017.
- [23] Wen X, Gong F, Fang Z, et al. Queue Length Algorithm Based on Electric Police Vehicle Passing Data [J]. Heilongjiang Transportation Science and Technology, 2018, 41(05): 171-172.
- [24] Luo X, Ma D, Jin S, et al. Queue Length Estimation for Signalized Inter sections Using License Plate Recognition Data[J]. IEEE Intelligent Transportation Systems Magazine, 2019, 11(3): 209-220.
- [25] Ju J. Research on Queue Length of Intersection Traffic Flow Based on Electronic Police Data [D]. Shanghai University of Applied Technology, 2021.
- [26] Le H. Maximum Queue Length Estimation at Signalized Intersections Based on License Plate Data [D]. Southwest Jiaotong University, 2021.
- [27] Zhang W, Ye J, Bai Z, et al. Intersection Queue Length Estimation Method Based on Connected Vehicle Trajectory Data [J]. China Journal of Highway and Transport, 2022, 35(3): 216-225.
- [28] Wei L, Li J, Xu L, et al. Queue length estimation for signalized intersections under partially connected vehicle environment[J]. Journal of Advanced Transportation, 2022(5): 1-11.
- [29] Wang Q. Research and Application of Timing Optimization Scheme for Short-Term Traffic Flow at Urban Intersections [D]. University of Science and Technology of China, 2019.
- [30] Qi X. Research on Dynamic Timing Algorithm of Traffic Lights Based on Improved PSO [D]. Northeast Petroleum University, 2021.
- [31] Xu H, Liu X, Lin S, et al. Research on Optimization of Intersection Signal Control Based on Traffic Flow Forecasting[J]. Xian, China: American Society of Civil Engineers (ASCE), 2021, 37(06): 350-361.
- [32] Zhao Y. Research on Intersection Signal Optimization Method Considering Low Emissions [D]. Jilin University, 2022.
- [33] Wu Y. Video-Based Intersection Traffic Flow Detection and Signal Timing Optimization Algorithm [D]. Qingdao University of Technology, 2022.
- [34] Zhao W, Ye Y, Ding J, et al. IPDALight: Intensity- and phase duration-aware traffic signal control based on Reinforcement Learning[J]. Journal of systems architecture, 2022, 12(01): 123.
- [35] Hao Q, Weishi Z. Short-Term Traffic Flow Prediction and Signal Timing Optimization Based on Deep Learning [J]. Wireless Communications and Mobile Computing, 2022, (23): 20.