Research Article

An Adaptive Model for Calculating the Correlation Degree of Multiple Adjacent Signalized Intersections

Linhong Wang[1] and Yiming Bie[2]

[1] *College of Transportation, Jilin University, Changchun 130022, China*
[2] *School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150091, China*

Correspondence should be addressed to Yiming Bie; yimingbie@126.com

As an important component of the urban adaptive traffic control system, subarea partition algorithm divides the road network into some small subareas and then determines the optimal signal control mode for each signalized intersection. Correlation model is the core of subarea partition algorithm because it can quantify the correlation degree of adjacent signalized intersections and decides whether these intersections can be grouped into one subarea. In most cases, there are more than two intersections in one subarea. However, current researches only focus on the correlation model for two adjacent intersections. The objective of this study is to develop a model which can calculate the correlation degree of multiple intersections adaptively. The cycle lengths, link lengths, number of intersections, and path flow between upstream and downstream coordinated phases were selected as the contributing factors of the correlation model. Their jointly impacts on the performance of the coordinated control mode relative to the isolated control mode were further studied using numerical experiments. The paper then proposed a correlation index (CI) as an alternative to relative performance. The relationship between CI and the four contributing factors was established in order to predict the correlation, which determined whether adjacent intersections could be partitioned into one subarea. A value of 0 was set as the threshold of CI. If CI was larger than 0, multiple intersections could be partitioned into one subarea; otherwise, they should be separated. Finally, case studies were conducted in a real-life signalized network to evaluate the performance of the model. The results show that the CI simulates the relative performance well and could be a reliable index for subarea partition.

1. Introduction

When the network-wide traffic signal control strategy is implemented to an urban road network, subarea partition must be conducted to divide the network into some small subareas. The subarea can be classified into three types. The first type only includes one signalized intersection. The second type includes several intersections located on one artery. The third type includes several intersections located on some intersected arteries. For the three type subareas, isolated signal control mode, arterial signal coordination mode, and area signal coordination mode should be implemented to the intersection(s), respectively. Accordingly, subarea partition is the basis for network-wide signal control. A well subarea partition algorithm can improve the adaptive level of the signal control system significantly [1–3].

In fact, the most important part of subarea partition is calculating the correlation degree among adjacent signalized intersections. The purpose of subarea partition is to improve the whole control performance of the network. If the performance could be improved, then adjacent intersections can be partitioned into one subarea. Otherwise they should be partitioned into different subareas. Thus, a straightforward method for subarea partition is to calculate and compare the control performances under the coordinated control mode and the isolated control mode. However, this method is not easy to implement in practice although it is theoretically straightforward. This is because it would be time consuming to iteratively estimate the performances of hundreds of adjacent intersections in the network. Thus, an effective alternative is to propose an index that reflects the changes in control performance indirectly; this is called the correlation index (CI) in this paper.

Because of the importance of subarea partition, the related researches have been conducted from the 1970s [4–11]. Yagoda et al. defined a coupling index (I) as $I = V/L$.

A higher value of the index would indicate that it would be preferable to have the two intersections coupled. V is the traffic volume and L is the link length between two adjacent intersections [12]. Chang brought forward the interconnection desirability index by considering both flow pattern and platoon dispersion characteristic. If the value of the index was greater than 0.35, the two intersections should be grouped together [13]. Moore and Jovanis [14] thought that adjacent intersections can be partitioned together when their v/c ratios were close. However, the threshold of v/c ratios was not determined. Lin and Huang [15] established a simple model to identify whether two adjacent signalized intersections can be coordinated based on the block length. The block length is affected by original platoon size and platoon completeness ratio. However, the block length cannot depict the impacts of all contributing factors on subarea partition. Lu et al. [16] selected link length, link volume, and cycle length as contributing factors and then developed two models to describe the correlation between two adjacent intersections and between multiple adjacent intersections, respectively, according to expert knowledge. Bie et al. [17] chose not only cycle length but also platoon length as contributing factors. An integrated correlation model was developed to quantify the correlation between two adjacent intersections. However, the determination of the weighted coefficients in the model and the threshold value were dependent on past experience and lacked a theoretical basis.

As described above, some valuable results have been achieved on the subject of correlation model development. However, there are still some deficiencies that needed to be resolved, which are summarized as follows.

(1) The contributing factors to the correlation index and the model structure are mainly determined according to the past experience of traffic engineers. Most weighting factors and parameters are fixed and cannot vary as the traffic state changes. Thus, the developed models cannot quantify the correlation between adjacent intersections accurately.

(2) The threshold value for the correlation has not been theoretically justified. The threshold value directly determines whether adjacent intersections can be partitioned into one subarea. However, in existing studies, the threshold value is again determined based on the experience of engineers. Accordingly, the subarea partition results may not be optimal.

(3) The existing studies have not yet developed a model that can measure the correlation among multiple intersections. In most cases, one subarea includes more than two intersections. Thus, it is inevitable to calculate the correlation among multiple intersections. However, existing studies only calculate the correlation between two adjacent intersections.

Targeting the real-time traffic control systems, an appropriate correlation index should reflect the dynamic nature of traffic flow and interactions among adjacent intersections. In this paper, we propose an approach that can estimate the dynamic correlation among multiple adjacent intersections for the purpose of subarea partition. The key innovation of this study is that the contributing factors are expressed as variables and the relation model between correlation and the variables is developed. The correlation can be calculated dynamically as the traffic state changes. Moreover, our modeling approach provides a reasonable way to determine the threshold value of correlation at which adjacent intersections should be grouped into one subarea.

2. Contributing Factors for Subarea Partition

The determination of contributing factors for subarea partition is the basis for developing the correlation model. Correlation model is the core of subarea partition, thus the principles of subarea partition must be obeyed when developing the model. If all intersections in the urban network execute isolated signal control mode, it is unnecessary to develop the correlation model because isolated control mode only concerns the intersection itself. However, when we identify whether multiple adjacent intersections can be partitioned together to execute signal coordination mode, the factors that may affect signal coordination should be analyzed. Thus, the factors that can improve signal coordination performance are also those that affect subarea partition. In this paper, the following four factors are selected.

(1) Difference between Cycle Lengths of Adjacent Intersections. When intersections in a subarea execute signal coordination mode, they should run the common cycle length (CCL). The CCL can help the signal controller to achieve maximum green wave bandwidth and maintain optimal offsets. Usually CCL equals the maximum cycle length (MCL) of all optimal cycle lengths (OCL) of the intersections. OCL is the cycle length of one intersection when it executes isolated control mode. However, the implement of CCL would increase the vehicle delay, because the CCL is larger than the OCLs. The increased delay is directly proportional to the difference between the CCL and the OCL. If the difference is large enough, the increased vehicle delay may be larger than the reduction in delay achieved by the signal coordination.

(2) Link Length between Two Adjacent Intersections. Signal coordination can reduce vehicle delay by providing green time for the approaching traffic platoon. A compact platoon is proven to be effective in improving coordination benefit because more vehicles can pass through the stop line during green time. According to Robertson's platoon dispersion model [18], the original platoon that departs from the upstream stop line would disperse to some extent and the dispersion level is directly proportional to the link length. Accordingly, a long link would pay a negative impact on the arterial progression.

(3) Path Flow between Upstream and Downstream Coordinated Phases. When a platoon travels to the downstream intersection, some vehicles in the platoon may turn to uncoordinated downstream phases. Therefore, the path flow between the two coordinated phases would be less than the original flow that departs from the upstream coordinated

phase. A greater path flow indicates that more vehicles are traveling between the two phases and there is greater opportunity to reduce vehicle delays through signal coordination. Thus, the coordination benefit is directly proportional to the path flow.

(4) Number of Intersections. According to the field experiences of engineers, the control performance of signal coordination would decrease with the increase in the number of intersections in the subarea. In fact, the reason for performance decrease is due to the cumulative negative impacts incurred by the other three contributing factors. In reality, there may be large differences between the cycle lengths of adjacent intersections, long link length, and low path flow. Therefore, from this perspective, the number of intersections is not a direct contributing factor to subarea partition, but an indirect one caused by the other three factors.

However, if we assume there is no difference between the cycle lengths of adjacent intersections, the link lengths are exactly right, and the path flows are great enough, then these three factors would not have negative impact on signal coordination performance. In such situation, the larger the number of intersections in the subareas, the better the coordination performance would be. Therefore, from this perspective, the number of intersections would be a direct contributing factor to subarea partition.

3. Modeling the Correlation

3.1. Performance Index. Let us take two adjacent signalized intersections i and j as an example. Their total vehicle delays per cycle are denoted by D_i and D_j when they are running the isolated control mode. D_i and D_j can be obtained using the delay function from HCM 2000 [19]. If the two intersections are grouped into one subarea, they will run signal coordinated mode. In this case, the total vehicle delays per cycle are denoted by $\widehat{D}_i$ and $\widehat{D}_j$. Therefore, we propose an index PI, and it indicates the improved performance achieved by grouping i and j into one subarea and can be calculated using

$$\begin{aligned} \text{PI} &= D - \widehat{D} \\ &= \left(D_i \frac{3600}{C_i} + D_j \frac{3600}{C_j} \right) - \left(\widehat{D}_i + \widehat{D}_j \right) \frac{3600}{\max\left(C_i, C_j \right)}. \end{aligned} \quad (1)$$

In (1), C_i is the OCL of i when it is run in the isolated control mode. $\max(C_i, C_j)$ is the CCL when the signal coordinated mode is run, D is the total vehicle delay per hour under the isolated control mode, and $\widehat{D}$ is the total vehicle delay per hour under the coordinated mode. All delays are unified into the total vehicle delay per hour (the unit of PI is s/h). Because the OCL of a nonseed intersection may be smaller than the CCL, the number of arriving vehicles in each cycle may also be smaller than that when the signal coordinated mode is operated.

If PI is larger than 0, then the signal coordinated mode achieves better control performance than the isolated control

mode; accordingly i and j can be grouped into one subarea. Otherwise, they should be assigned to different subareas.

The performance of the signal coordination is also affected by the coordination type (i.e., one-directional or two-directional coordination); different coordination types will produce different partition results. In this paper, we only study the relations between the correlation index and the contributing factors under two-directional coordination conditions, which is useful to judge whether adjacent intersections can be partitioned into one subarea to execute two-directional coordination algorithm.

The best way to analyze the impact of a contributing factor on PI is to determine the relationship between them based on theoretical derivation and then calculate the change in PI resulting from a unit change in the contributing factor. However, the relationship is difficult to determine due to the complex nature of vehicle movements between consecutive intersections. There is no universal function that can be used to calculate $\widehat{D}$. Thus, numerical experiments are used for this paper. Experimental data are obtained relating the PI to the contributing factors, and then relation models between CI and the contributing factors are fitted in Sections 3.2 to 3.5.

An important issue that should be noted is that the traffic control objective is closely related to the traffic state of the intersection. For example, the control objective for an oversaturated intersection is usually set as minimizing the queue length to avoid queue spillovers, while that of an unsaturated intersection is usually set as minimizing vehicle delay. In this paper, we focus on unsaturated intersections and minimizing vehicle delay is selected as the control objective. Thus, in (1) PI indicates the reduced vehicle delay when using the signal coordinated mode instead of the isolated control mode.

Differences in cycle lengths only affect the control performances of the seed intersection and nonseed intersection. The link length and path flow only affect the performances of two adjacent intersections. Therefore, a network composing of two intersections is selected to conduct numerical experiments on these three factors.

3.2. Impact of the Difference in Cycle Lengths on Correlation. Again taking intersections i and j as an example, let the link length L equal 400 m and let the average vehicle speed be 12 m/s. j is the seed intersection and i is the nonseed intersection. Thus, when run in coordinated mode, the OCL of j is set as the CCL. The difference between the OCL of i and the CCL is increased using a step of 5 s, while the link length and path flow are fixed. Table 1 shows the results for two experimental scenarios, where the CCL equals 125 s and 110 s, respectively. In Table 1, the improvement is shown by the ratio of PI to D. The larger is the ratio, the better are the signal coordination benefits.

For example, in scenario 1 when the OCL of i equals 90 s, the difference between the cycle lengths is 35 s and PI equals 20,044 s. The improvement ratio is only 5.7%. However, when the OCL of i increases to 125 s, the improvement ratio is 17.24%. The difference between cycle lengths is negatively correlated to the improvement ratio and a similar trend

TABLE 1: Results of the experimental scenarios.

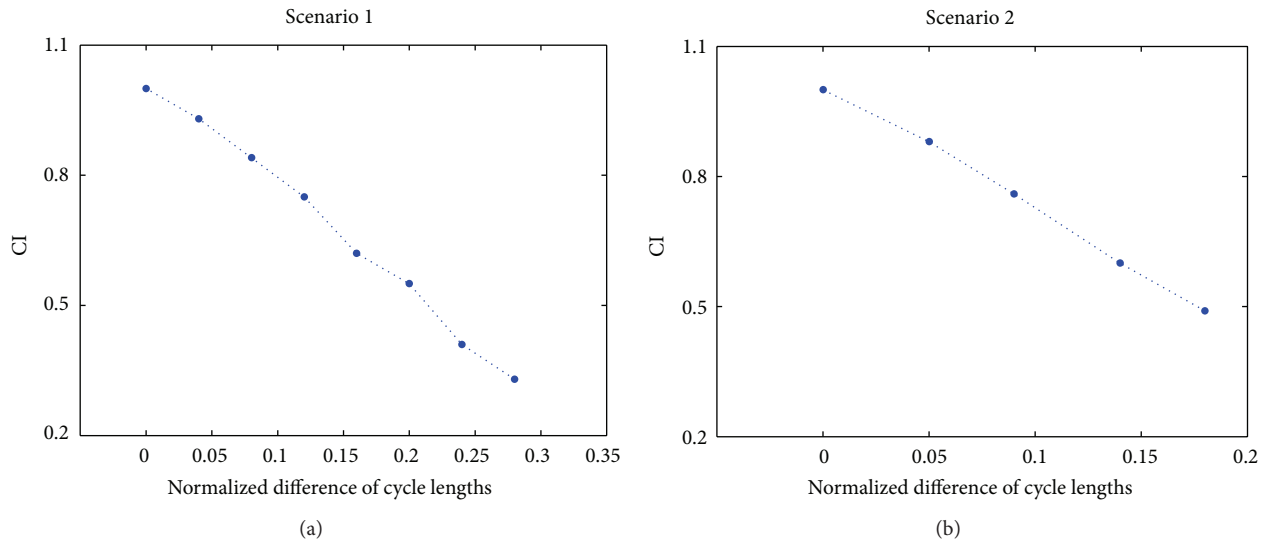| Scenario	CCL/s	OCL of i/s	D/s·h^{-1}	PI/s·h^{-1}	Improvement ratio
1	125	90	351576	20044	5.70%
		95	358915	25438	7.09%
		100	371277	34919	9.41%
		105	378642	40453	10.68%
		110	390619	50311	12.88%
		115	399471	57848	14.48%
		120	408876	65726	16.07%
		125	416925	71878	17.24%
2	110	90	322504	22407	6.95%
		95	331007	28399	8.58%
		100	342658	37425	10.92%
		105	351739	44573	12.67%
		110	360691	51658	14.32%

(a)

(b)

FIGURE 1: Scatter diagrams between CI and the normalized difference in cycle lengths.

is found in other scenarios. The results indicate that the difference between cycle lengths has a substantial impact on PI.

The data shown in Table 1 can also be interpreted as follows. When the difference equals 0, signal coordination achieves the maximum benefits, which means that the correlation between the two intersections is maximal and the correlation index (CI) is 1.0. However, when the difference equals 35 s, the ratio of the improvement to that when the difference equals 0 is only 33%; that is, CI decreases to 0.33.

As can be seen from Table 1, CI is affected not only by the difference in cycle lengths, but also by the CCL. Thus, the difference in cycle lengths is normalized as follows:

$$C_D = \frac{C_j - C_i}{C_j}. \tag{2}$$

Scatter diagrams between CI and C_D are shown for the two scenarios in Figure 1.

There is a strong linear relationship between CI and C_D, and a linear function can be used to model their relationship. The linear functions for the two scenarios are as follows:

$$CI = \begin{cases} -2.471C_D + 1.024, & R^2 = 0.995, \\ -2.894C_D + 1.009, & R^2 = 0.996. \end{cases} \tag{3}$$

Linear functions can fit the scatter diagrams very well, as indicated by the high R-square values. However, the problem is that the slopes of the two lines are different from each other. This indicates that, in different scenarios, C_D will have a different impact on CI. Therefore, a universal function is needed that depicts the relationship between the two variables, in which the slope varies as the traffic state changes. The linear functions in (3) can be replaced by the following format:

$$CI = \alpha_1 C_D + b_1, \tag{4}$$

where α_1 is the slope of the function, and b_1 is the intercept.
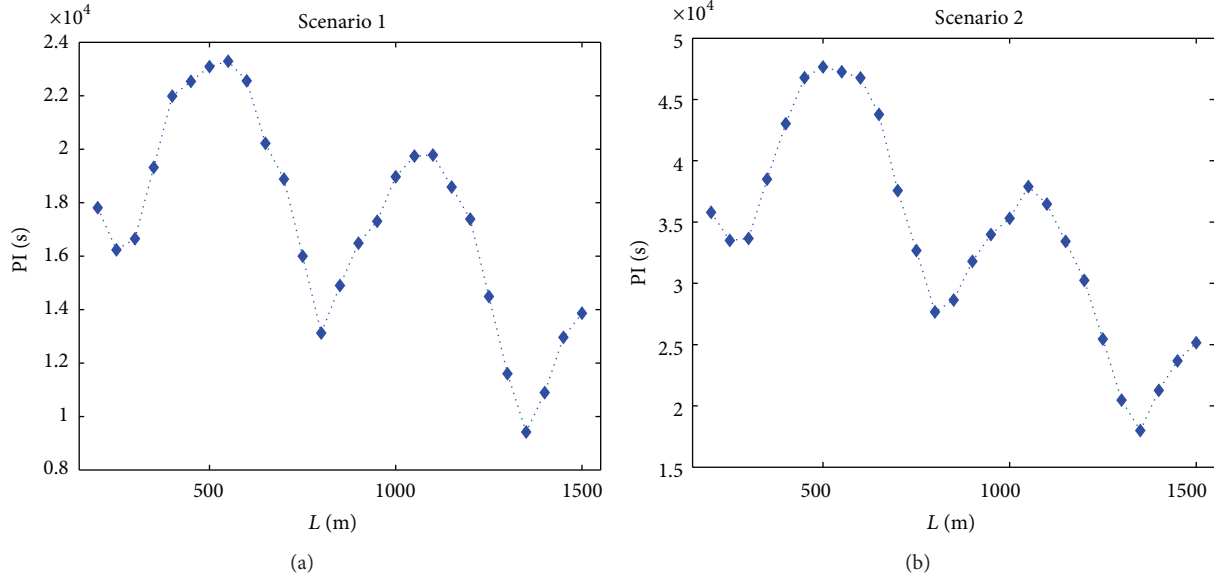
FIGURE 2: Scatter diagrams between L and PI for the two scenarios.

Theoretically, b_1 should equal 1. However, due to fitting errors in (3), marginal differences exist between the intercepts and 1. In this study, however, the differences are neglected and the value of b_1 is simply set to 1. Therefore, only α_1 needs to be fitted.

α_1 represents the change in CI resulting from a unit change in C_D. It may be affected by many factors, such as the green split and the volume to capacity ratio (v/c) of coordinated phases, the v/c of uncoordinated phases, the total v/c, the CCL and the intersection saturation degree x. To study the impacts of these factors on α_1, a multivariate regression can be used. First, the factors are introduced into the fitting function and then the variables are eliminated stepwise according to the correlation until the required precision is achieved. Then, the residual variables in the function are those with a significant impact on α_1. The amount of data in Table 1 is not sufficient to carry out a regression. Accordingly, another 10 experiments are carried out, providing a total of 12 groups of data. The resulting regression function is shown in

$$\alpha_1 = 14.916 - 53.963x + 4.831\lambda_c + 36.281Y,$$

$$R^2 = 0.920.$$

(5)

In the function, λ_c is the green split of the coordinated phase and Y is the total v/c. Only these three are retained; the other variables are excluded. An F test is carried out to identify whether the real value and fitted value of α_1 differ significantly (under a significance level of 0.05). The results of the test are acceptable.

3.3. Impact of Link Length on Correlation. Numerical experiments are conducted to establish the relationship between the link length L and the CI. During the experiments, L is increased from 200 m to 1500 m with a step of 50 m while the cycle lengths of the two intersections and the path

flow between coordinated phases are fixed in each scenario. Figure 2 shows the results for two typical scenarios. The cycle lengths of intersection i and j in the two scenarios are set to the same value, 90 s.

In Figure 2, L and PI have a curvilinear relationship and the two curve shapes are similar to each other. As the value of L increases, the value of PI increases up to a maximum, which is attained when L equals about $0.5C_jV$. As L increases further, PI decreases and reaches a minimum when L equals about $0.75C_jV$. The cycle is repeated as L increases with a step size of $0.5C_jV$ meters. The difference is that the values of the maximum and minimum decrease step by step. This changing tendency is significantly different to our usual view, because traditionally we would assume that PI would decrease linearly with an increase in L. However, the traditional view is based on one-directional signal coordination, and in this study we mainly focus on the dual-directional signal coordination of two adjacent intersections. When L equals an integer multiple of $0.5C_jV$, dual-directional interactive coordination is the best approach and thus the traffic control can achieve the maximum benefits. When L equals an odd multiple of $0.25C_jV$, dual-directional coordination produces the worst benefit and thus PI is minimal.

In numerical experiments, L is increased with a step of 50 m; however, the integer multiples of $0.25C_jV$ are not equal to the integer multiples of 50. Thus, to obtain the exact curvilinear relationship between L and PI, the values of PI corresponding to integer multiples of $0.25C_jV$ are included.

When PI achieves its maximum at L equals $0.5C_jV$, grouping the two intersections into one subarea achieves the maximum control benefits; in this case, CI equals 1. In other cases, CI equals the ratio of PI to the maximum value of PI. L is normalized as follows:

$$L_c = \frac{L}{L_{\max}},$$

(6)

where L_c is the normalized link length.

Scenario	First minimum	First maximum	Second minimum	Second maximum	Third minimum
1	0.663	1	0.533	0.858	0.4
2	0.674	1	0.527	0.816	0.372
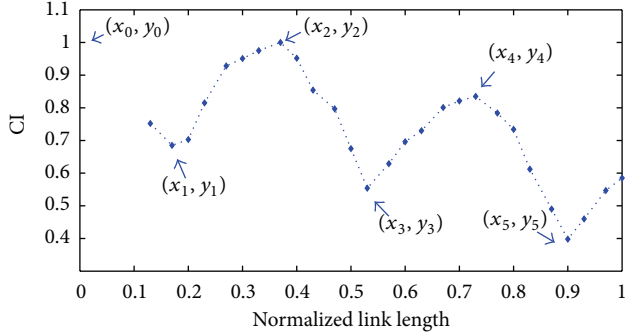
FIGURE 3: Scatter diagram between L_c and CI (scenario 1).

L_{max} is set as 1500 m in this study because signal coordination is not suitable when L is larger than 1500 m according to engineering experience. The relationship between CI and L_c is similar to that shown in Figure 2. Thus, the scatter diagrams are not displayed for these two variables.

Following detailed analysis, it was concluded that a piecewise linear function may be suitable to fit the curve. Moreover, the differences between adjacent minimal values of CI are nearly constant, while the differences between each maximal value of CI and its previous minimal value are also nearly constant. For scenarios 1 and 2, the extreme values of CI corresponding to integer multiples of $0.25C_jV$ are shown in Table 2.

In scenario 1, the difference between the first maximum and the first minimum equals 0.337. The difference between the second maximum and the second minimum equals 0.325. The two numbers are very close. The difference between the first minimum and the second minimum equals 0.130, while the difference between the second and third minimum equals 0.133. These numbers are also very close. In scenario 2, the same phenomenon is observed. The extreme points of the L_c-CI curve are labeled and shown in Figure 3.

In the numerical experiments, L ranges from 200 to 1500. However, theoretically, it is only known that L should be larger than 0. Thus, in Figure 3, the curve is extended to the interval $[0, 200]$. However, the values of CI corresponding to that interval do not have any practical meaning because, when L is smaller than 200 m, the objective of the signal control should be to minimize the queue length or avoid queue spillovers, rather than to reduce vehicle delay. On the basis of the changing tendency of CI, y_0 should be larger than 1. But again this has no practical meaning, so it is set equal to 1 for convenience.

To fit the curve in Figure 3 using a piecewise linear function, the values of y_1 and y_3 are needed. The difference

between y_2 and y_1 represents the reduction in CI when the link length decreases from $0.5C_jV$ to $0.25C_jV$. Similarly, the difference between y_2 and y_3 represents the reduction in CI when the link length increases from $0.5C_jV$ to $0.75C_jV$. The size of the reduction is affected by many factors, such as the CCL, the green split and the v/c of the coordinated phase, the v/c of the uncoordinated phases, the total v/c, and the saturation degree, x. The multivariate regression method described in Section 3.2 is also adopted here to determine which of the above factors have a significant impact on y_1 and y_3. In total, 10 groups of data are used in the regression.

The regression functions are shown in (7). In the two functions, only x, y_c, and Y are retained and the other variables are excluded. y_c is the v/c ratio of the coordinated phase:

$$
\begin{aligned}
y_1 &= 0.381 + 0.992 \cdot x - 0.870 \cdot Y \\
&\quad + 0.291 \cdot y_c, \\
R^2 &= 0.950, \\
y_3 &= -2.716 + 11.682 \cdot x - 9.814 \cdot Y \\
&\quad + 1.942 \cdot y_c, \\
R^2 &= 0.960.
\end{aligned}
\tag{7}
$$

Then, y_n in the L_c CI curve equals

$$
y_n = \begin{cases}
1.0, & n = 0, \\
0.381 + 0.992 \cdot x - 0.87 \\
\quad \cdot Y + 0.291 \cdot y_c, & n = 1, \\
1.0, & n = 2, \\
-2.716 + 11.682 \cdot x - 9.814 \\
\quad \cdot Y + 1.942 \cdot y_c, & n = 3, \\
y_{n-1} + y_{n-2} - y_{n-3}, & n \geq 4 \text{ and even integer}, \\
y_{n-2} + y_{n-4} - y_{n-2}, & n \geq 4 \text{ and odd integer}.
\end{cases}
\tag{8}
$$

The nth one in the sets of lines can be fitted by

$$
\text{CI}(n) = \alpha_2 \cdot L_c + b_2,
$$

$$
\alpha_2 = \frac{y_n - y_{n-1}}{x_n - x_{n-1}}, \quad b_2 = \left(y_{n-1} - \frac{y_n - y_{n-1}}{x_n - x_{n-1}} \cdot x_{n-1} \right),
\tag{9}
$$

where α_2 is the slope of the line, and b_2 is the intercept.

TABLE 3: Impact of the difference between q_s and $q_{s\,\max}$ on control benefits.

Scenario	P_q	D_u/s	D_c/s	P_{dk}/s	P_{du}/s	PI_n	P_n
	0%	265103	77701	0	0	0	0.00%
	5%	271860	72611	5090	6757	1667	3.08%
	10%	276639	69391	8310	11536	3226	5.95%
	15%	283045	67884	9817	17942	8124	14.99%
1	20%	287528	64228	13473	22425	8951	16.52%
	25%	297323	59695	18006	32220	14214	26.23%
	30%	299637	57025	20676	34534	13858	25.57%
	35%	305460	53310	24391	40357	15965	29.46%
	40%	309357	49541	28160	44254	16094	29.69%
	0%	198840	59700	0	0	0	0.00%
	5%	202996	58878	613	4156	3543	5.95%
	10%	206948	57892	3032	8107	5075	8.52%
	15%	210717	56741	5926	11877	5950	9.99%
2	20%	212268	56461	7176	13428	6253	10.49%
	25%	215648	54948	9405	16808	7403	12.43%
	30%	218853	53261	11658	20013	8354	14.02%
	35%	221884	51394	12925	23044	10118	16.98%
	40%	224744	49342	14916	25903	10988	18.44%

For scenario 1, the function between L_c and CI is as follows:

$$\text{CI} = \begin{cases} -1.872 \cdot L_c + 1.0, & 0 \le L_c \le 0.18, \\ 1.872 \cdot L_c + 0.326, & 0.18 < L_c \le 0.36, \\ -2.594 \cdot L_c + 1.934, & 0.36 < L_c \le 0.54, \\ 1.871 \cdot L_c - 0.478, & 0.52 < L_c \le 0.72, \\ -2.595 \cdot L_c + 2.738, & 0.72 < L_c \le 0.90, \\ 1.871 \cdot L_c - 1.282, & 0.90 < L_c \le 1.0. \end{cases} \quad (10)$$

The results of the F test show that there is no significant difference between the real values and the fitted values of CI. Thus, (10) can be used to depict the relationship between L_c and CI in scenario 1.

3.4. Impact of Path Flow on Correlation. Let q_s denote the path flow. The maximum value of q_s (denoted by $q_{s\,\max}$) equals the maximum flow of the upstream coordinated phase. Since signal coordination is not suitable to be implemented to the near saturated intersection, $q_{s\,\max}$ is taken to be the maximum historical traffic volume of the upstream coordinated phase when the saturation degree is no larger than x_c. In this paper, it is recommended that x_c be set at 0.90.

Again taking intersections i and j as an example, two scenarios are tested. Intersections i and j have the same cycle length and saturation degree. The $q_{s\,\max}$ values in the two scenarios are 640 pcu/h and 670 pcu/h, respectively. The path flows in each direction are the same. q_s is decreasing from $q_{s\,\max}$ with a decrement of 5%, while the cycle length and link length remain fixed.

Table 3 shows the results. In the table, P_q is the ratio of $(q_{s\,\max} - q_s)$ to $q_{s\,\max}$, which increases from 0 to 40%. D_u is the total vehicle delay per hour of uncoordinated phases. To keep the cycle length fixed in each scenario, the reduced traffic flow is transferred equally among the uncoordinated phases. Thus, the traffic flow and vehicle delay in the uncoordinated phases will increase. D_c is the total vehicle delay per hour of coordinated phase. Because the traffic volume of the coordinated phase decreases gradually, D_c also decreases. P_{dk} is the decreased vehicle delay of coordinated phase compared with that when P_q equals 0. P_{du} is the increased vehicle delay of uncoordinated phases compared to that when P_q equals 0. PI_n is the negative performance index resulting from the decrease in q_s. P_n is the ratio of PI_n to the PI when P_q equals 0.

In Table 3, the signal control between the two intersections achieves maximum benefits when P_q equals 0. Therefore, CI equals 1. P_n indicates the degradation of PI due to an increase in P_q. Taking scenario 1 as an example, P_n is 14.99% when P_q equals 15%. In this case, the control benefits equal 85.01% of the PI when P_q equals 0; that is, CI equals 0.8501. CI is the difference between 1 and P_n. P_q can be used as the normalized index of path flow. The scatter diagrams between P_q and CI for the two scenarios are shown in Figure 4.

As can be seen from Figure 4, there is a strong linear relationship between P_q and CI, so a linear function can be fitted to the scatter diagrams. The fitted functions are shown in (11). The R squares of the two functions are both larger than 0.94, which indicates that a linear function fits the relationship well:

$$\text{CI} = \begin{cases} -0.828 \cdot P_q + 0.997, & R^2 = 0.950, \\ -0.401 \cdot P_q + 0.973, & R^2 = 0.940. \end{cases} \quad (11)$$

The problem is that, as in the earlier case, the slopes of the two functions are different. The slope may again be affected by our previously used set of factors, namely the CCL, the green split and the v/c of the coordinated phase, the v/c of the
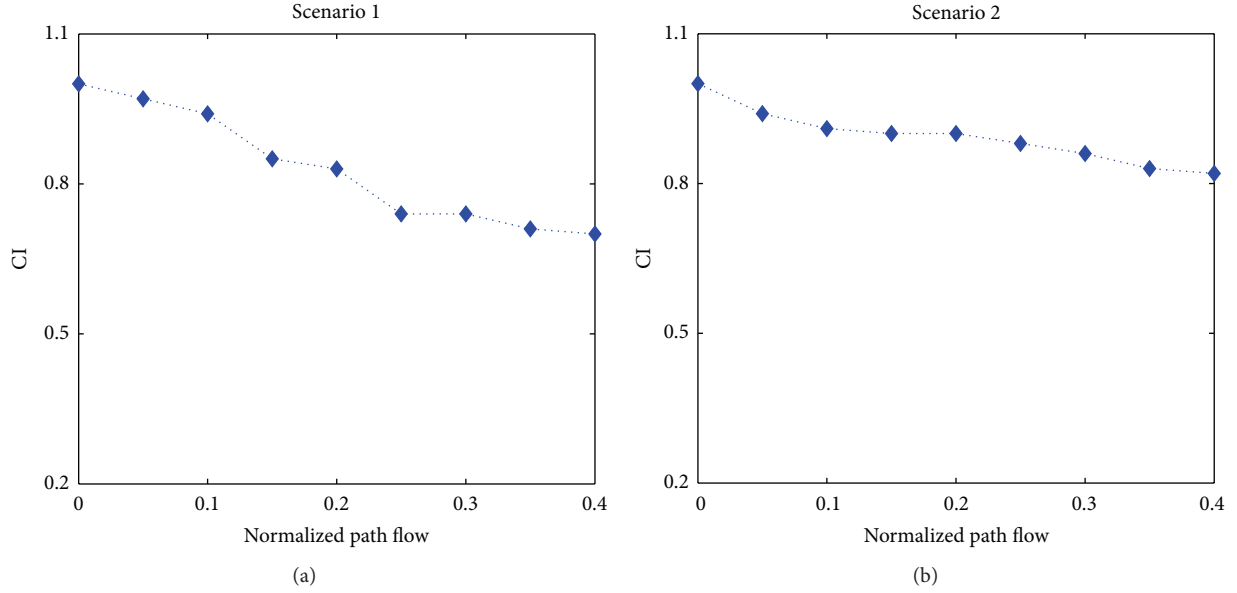
Scenario 1 (a)

Scenario 2 (b)

Figure 4: Scatter diagrams between normalized path flow and CI.

uncoordinated phases, the total v/c, and the saturation degree x. The multivariate regression method shown in Section 3.2 is again adopted here to distinguish whether any of these factors has a significant impact on the slope. The universal function between P_q and CI may be written as

$$\text{CI} = \alpha_3 \cdot P_q + b_3, \tag{12}$$

where α_3 is the slope and b_3 is the intercept.

As in the earlier case, b_3 equals 1 theoretically but there are marginal differences between the intercepts in (11) and 1 because of fitting error. Once again, in this study, the differences are neglected and b_3 is set as 1.

A total of 10 groups of data are used in the regression and the result is shown in

$$e^{\alpha_3} = 5.946 - 2.664 \cdot \log{(\text{CCL})} + 0.140 \cdot y_u$$
$$- 0.049 \cdot \lambda_c, \tag{13}$$

$$R^2 = 0.93.$$

α_3 is affected by the CCL, y_u, and λ_c. Equation (13) can be rewritten as

$$\alpha_3 = \ln{(5.946 - 2.664 \cdot \log{(\text{CCL})}}$$
$$+ 0.140 \cdot y_u - 0.049 \cdot \lambda_c). \tag{14}$$

In the above numerical experiments, the path flows in the two directions are assumed to be the same. However, in reality they may be different. Further numerical experiments are conducted and show that P_q in (12) can be calculated as follows:

$$P_q = \frac{\left(P_{q1} + P_{q2}\right)}{2}, \tag{15}$$

where P_{q1} and P_{q2} are the ratios between the current path flow and the maximum historical path flow in each direction.
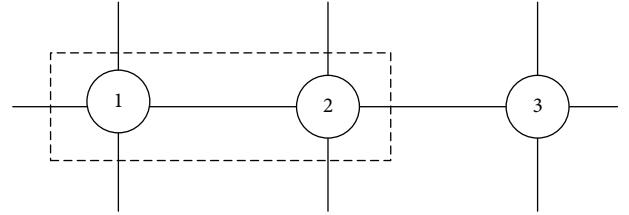

Figure 5: Sketch of three adjacent signalized intersections.

3.5. Impact of Number of Intersections on Correlation. The number of intersections in the subarea affects the number of coordinated traffic streams. As was shown in Figure 5, the arrival streams at intersection 1 in the direction of west to east and that at intersection 3 in the direction of east to west cannot be coordinated because they arrive from outside of the subarea. The ratio of coordinated streams to total streams of coordinated phases in the network is 2/3. With an increase in the number of intersections, the ratio will decrease. Thus, the proportion of vehicles that can be coordinated will increase and the coordinated control performance will increase. To validate the above deduction, numerical experiments are conducted.

In the numerical experiments, the values of the three other contributing factors are fixed. The difference in cycle lengths is set to 0. The link lengths between adjacent intersections are $0.5C_cV$. The path flows of adjacent intersections are set equal to their maximum historical values. In scenarios 1 and 2, all intersections have four phases. In scenarios 3 and 4, all intersections have three phases. To study the impact of the number of intersections on the coordination performance, we simulate subareas with 2, 3, 4, 5, 6, and 7 intersections in each scenario. The results for the four scenarios are shown in Table 4.

TABLE 4: Results for the four experimental scenarios.

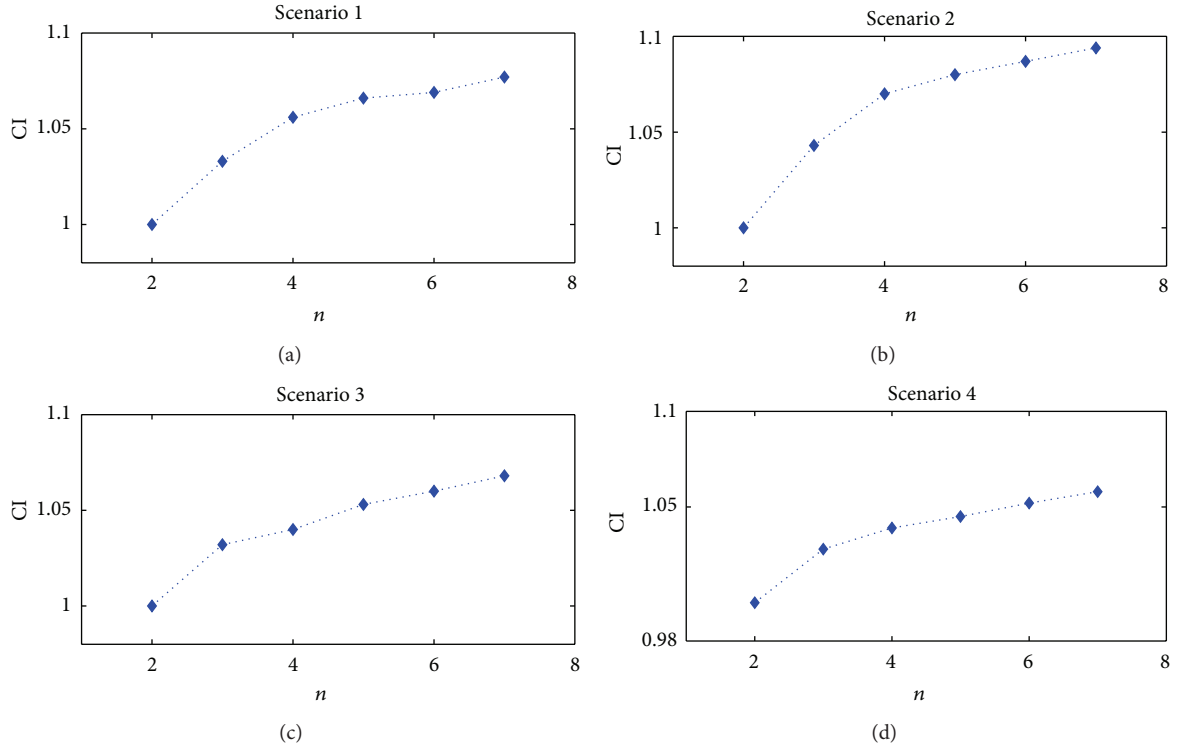|	Scenario 1		Scenario 2		Scenario 3		Scenario 4
n	Improvement ratio	n	Improvement ratio	n	Improvement ratio	n	Improvement ratio
2	8.81%	2	15.35%	2	13.70%	2	17.10%
3	12.35%	3	20.30%	3	17.39%	3	20.39%
4	14.92%	4	23.37%	4	18.21%	4	21.65%
5	15.96%	5	24.56%	5	19.76%	5	22.36%
6	16.34%	6	25.40%	6	20.55%	6	23.16%
7	17.20%	7	26.20%	7	21.38%	7	23.90%

(a)

(b)

(c)

(d)

FIGURE 6: Scatter diagram between number of intersections in the subarea and CI.

In Table 4, n is the number of intersections in the subarea. Taking scenario 1 as an example, when n equals 2, the improvement ratio is 8.81%. With an increase in n, the improvement ratio also increases. When n equals 7, the improvement ratio reaches 17.2%. This pattern can also be seen in the other scenarios, demonstrating that the number of intersections indeed impacts upon subarea partition.

Since there is no difference between the cycle lengths in each scenario, the lengths of the links are optimal and the path flows are equal to their maximum historical values. Therefore, when n equals 2, the value of CI for the two adjacent intersections is 1. However, the improvement ratio increases as n increases, so CI also increases. This means that CI will be larger than 1 when n is larger than 2. Based on the above analysis, CI can be calculated as follows:

$$\text{CI}(n) = \frac{1 + \text{PIC}(n)}{1 + \text{PIC}(2)}, \quad (16)$$

where PIC(n) is the improvement ratio when there are n intersections in the subarea.

The correlations between the intersections shown in Table 4 are calculated according to (16), and the scatter diagrams between n and CI for the four scenarios are shown in Figure 6.

From the figure, we can see that, with an increase in n, the increase in CI decreases. The logarithmic function can be used to fit the relation between n and CI. For the four scenarios, the fitting functions are as follows:

$$\text{CI} = \begin{cases} 0.061 \ln(n) + 0.963, & R^2 = 0.962, \text{ Scenario 1,} \\ 0.074 \ln(n) + 0.957, & R^2 = 0.960, \text{ Scenario 2,} \\ 0.052 \ln(n) + 0.969, & R^2 = 0.975, \text{ Scenario 3,} \\ 0.044 \ln(n) + 0.974, & R^2 = 0.968, \text{ Scenario 4.} \end{cases}$$
$$(17)$$

The R squares of the four functions are all larger than 0.96, which indicates that the logarithmic function fits the scatter

diagrams well. However, the slopes of the functions differ from each other, which demonstrates that the impact of the number of intersections on the correlation changes as the scenario changes. The slope is also affected by the same set of factors discussed in the previous sections (the CCL, the green split, etc.). The multivariate regression method shown in Section 3.2 is again adopted here to determine which if any of these have a significant impact on the slope. The universal function between n and CI is given by

$$\mathrm{CI} = \alpha_4 \cdot n + b_4, \tag{18}$$

where α_4 is the slope.

CI equals 1 when n is 2. Therefore, $b_4 = 1 - \alpha_4 \ln(2)$. Another 6 groups of data are obtained based on numerical experiments so that a total of 10 groups of data can be used to determine α_4. The outcome is as follows:

$$\alpha_4 = -0.263 - 0.511 y_c + 1.12 x - 0.255 \log(C_c). \tag{19}$$

The R square of the regression model is 0.952. α_4 is affected by three factors, the v/c of the coordinated phase, the saturation degree, and the CCL.

3.6. Integrated Correlation Model of Multiple Adjacent Intersections. In Sections 3.2 to 3.5, detailed relationships between various contributing factors and the CI are determined. To show their joint impact on the CI, the following integrated correlation model is proposed:

$$
\begin{aligned}
\mathrm{CI} &= \mathrm{CI}(n) \\
&\cdot \left\{ 1 - \left[\sum_{j=1}^{n} \left[\left(1 - \mathrm{CI}_{i,j}(C_D) \right) + \left(1 - \mathrm{CI}_{i,j}(L_c) \right) \right. \right. \right. \\
&\qquad \left. \left. \left. + \left(1 - \mathrm{CI}_{i,j}(P_q) \right) \right] \right] (n-1)^{-1} \right\}, \quad j \neq i, \\
&= \mathrm{CI}(n) \cdot \left\{ 1 - \left[\sum_{j=1}^{n} \left[3 - \left(\mathrm{CI}_{i,j}(C_D) + \mathrm{CI}_{i,j}(L_c) \right. \right. \right. \right. \\
&\qquad \left. \left. \left. \left. + \mathrm{CI}_{i,j}(P_q) \right) \right] \right] (n-1)^{-1} \right\},
\end{aligned}
\tag{20}
$$

where $\mathrm{CI}_{i,j}(C_D), \mathrm{CI}_{i,j}(L_s), \mathrm{CI}_{i,j}(P_q)$ and $\mathrm{CI}(n)$ can be obtained from (4), (9), (12), and (18), respectively.

In (20), i is the seed intersection of the subarea, and j is a nonseed intersection. The term $1 - (\mathrm{CI}_{i,j}(C_D))$ indicates the decrease in the correlation because of the difference between the cycle lengths of i and j. The terms in square brackets indicate the total decrease in the correlation due to adding j into the subarea. When j is not adjacent to i, $\mathrm{CI}_{i,j}(L_c)$ is the link length correlation between j and the intersection that is adjacent to j and located between j and i. Similarly, $\mathrm{CI}_{i,j}(L_c)$ is the path flow correlation between j and the intersection that is adjacent to j and located between j and i.

As analyzed above, the difference in cycle lengths, the path flow, and the link length only affect the correlation of two intersections. Therefore, the three related correlations are subtracted from 1. However, $\mathrm{CI}(n)$ is related to all of the intersections in the subarea. Thus, it is set as the base correlation of the subarea and multiplied by the other terms.

3.7. Threshold Value of CI. PI is the improved network performance when signal coordination is implemented. When PI is larger than 0, the coordination of adjacent intersections can obtain better performance than the isolated control mode. In such situations, the two can be grouped into the same subarea; otherwise, they must be placed in different subareas.

In this paper, the CI is developed in order to simulate PI. When PI is larger than 0, CI is also larger than 0. Therefore, 0 is set as the threshold value of CI, to determine whether adjacent intersections should be placed in the same subarea.

4. Case Study

In this section, field experiments are conducted to validate the correlation model developed in this paper. The experimental setup and then the results are described in the following subsections.

4.1. Experimental Setup. A small road network composed of four signalized intersections in Harberin, China, is selected as the experimental network. The four intersections are all located in the central area of the city. A sketch and phasing diagram of the network are shown in Figure 7. Songshan road is an arterial road. All intersections carry out a signal timing scheme with four phases. When executing the signal coordinated mode, the phases in the north-south direction are selected as the coordinated phases. The green lost time of each phase is 3 s and the saturation flow rate of each lane is 1700 pcu/h. For convenience, the four intersections are labeled intersections 1, 2, 3, and 4, respectively, from north to south. The link lengths between adjacent intersections are 730 m, 510 m, and 500 m.

An excellent CI should reflect the performance of the coordinated control mode relative to the isolated control mode precisely. Namely, the correlation should vary as the relative performance varies. Thus, five scenarios are tested. Because the signal coordinated mode is not suitable for those intersections facing oversaturated or close to oversaturated conditions, the five scenarios are all selected during off-peak hours. The duration of each scenario is 15 minutes. Table 5 displays the survey times of the five scenarios.

Traffic data are collected every 5 minutes and the timing schemes are optimized and updated every 15 minutes. Thus, the correlation among the four intersections is calculated once in each scenario. The calculation of the correlation requires some basic traffic parameters, such as cycle lengths, green splits of coordinated phases, and saturation degrees. The values of the parameters are shown in Table 6. C_i is the optimal cycle length of each intersection when operating in isolated control mode.
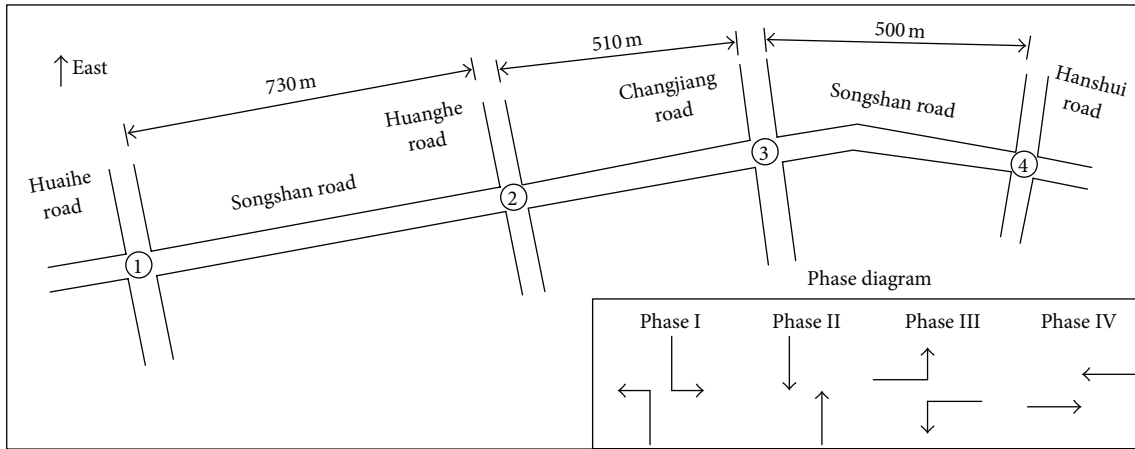
Figure 7: Sketch and phasing diagram of the investigated network.

Table 5: Survey times of the five experimental scenarios.

Scenario	Survey time
1	9:01 am~09:15 am
2	9:16 am~09:30 am
3	9:31 am~09:45 am
4	9:46 am~10:00 am
5	10:01 am~10:15 am

As can be seen from Table 6, the cycle length of intersection 2 is the longest in all scenarios. Thus, intersection 2 is set as the seed intersection when operating the coordinated mode.

4.2. Experimental Results and Analysis. Table 7 shows the results for the five scenarios.

From scenario 1 to scenario 5, the integrated correlations of the four intersections are 0.42, 0.38, 0.34, 0.51, and 0.55, respectively. Because the threshold value is 0, in the five scenarios, the four intersections can be grouped into one subarea. To test whether CI reflects PI accurately, we also give the values of PI and the improvement ratio of the signal coordination mode to the isolated control mode.

However, CI and the improvement ratio of PI have different dimensions and cannot be compared directly. Thus, the data for CI and the improvement ratio in Table 7 are normalized using a linear function. The normalized data are shown in Table 8.

The scatter diagrams for the scenarios between the normalized CIs and improvement ratios are shown in Figure 8.

From Figure 8 we can see that the two variables have similar changing tendency, but not exactly the same. This is because the correlation model was developed based on numerical experiments and there are some differences between the default values of the numerical experiments and the field values for the traffic parameters.

An F-test and a t-test are carried out to identify whether the two groups of normalized data differ significantly under a significance level of 0.95. The Data Analysis tool in Microsoft
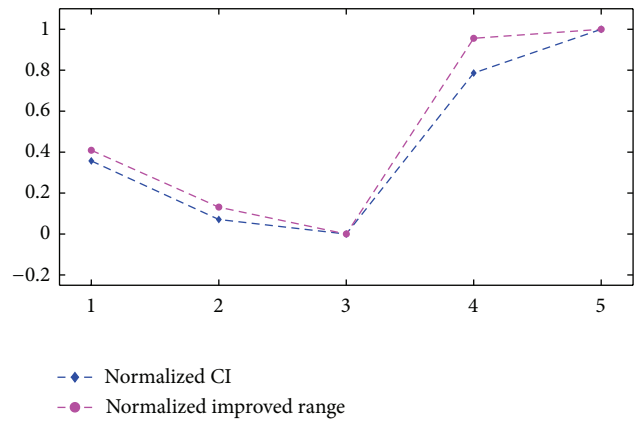

Figure 8: Comparison between normalized CI and improvement ratio.

Excel is convenient for executing these tests. The statistical results are shown in Table 9 and all are acceptable.

From the experimental results, we can see that the CI is closely related to the change in the PI, which indicates that the CI can be used as an alternative to the PI. Therefore, the correlation model developed in this paper is suitable for subarea partition.

5. Conclusions

This paper developed a correlation model for multiple adjacent intersections by taking into consideration four contributing factors: the difference between cycle lengths, the link length, the path flow between the upstream and downstream coordinated phases, and the number of intersections. A case study was used to explain the model and the results show that it is reliable for subarea partition.

In this paper, we selected four typical signal intersections located on an arterial road under moderately congested traffic conditions for our experiment. Although the traffic and signal timing conditions of the selected sites are pretty

TABLE 6: Parameters used to calculate the correlations among the intersections.

Scenarios	Intersections	C_i/s	λ_c	x	Y	Y_c	Y_u
Scenario 1	1	108	0.31	0.89	0.79	0.27	0.51
	2	114	0.32	0.89	0.80	0.29	0.51
	3	102	0.29	0.88	0.77	0.25	0.52
	4	106	0.27	0.88	0.78	0.24	0.54
Scenario 2	1	105	0.31	0.88	0.78	0.27	0.51
	2	113	0.33	0.89	0.80	0.29	0.50
	3	105	0.3	0.88	0.78	0.26	0.52
	4	97	0.28	0.87	0.76	0.24	0.52
Scenario 3	1	103	0.3	0.88	0.78	0.26	0.51
	2	110	0.29	0.89	0.79	0.26	0.53
	3	101	0.31	0.88	0.77	0.27	0.50
	4	95	0.29	0.87	0.76	0.25	0.51
Scenario 4	1	98	0.28	0.87	0.77	0.24	0.52
	2	101	0.3	0.88	0.77	0.26	0.51
	3	94	0.27	0.87	0.76	0.23	0.52
	4	92	0.31	0.86	0.75	0.27	0.48
Scenario 5	1	95	0.32	0.87	0.76	0.28	0.48
	2	99	0.33	0.87	0.77	0.29	0.48
	3	93	0.31	0.86	0.75	0.27	0.48
	4	88	0.31	0.86	0.74	0.27	0.47

TABLE 7: Results for the five experimental scenarios.

Parameters	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
$CI_{1,2}(C_D)$	0.86	0.81	0.82	0.91	0.89
$CI_{3,2}(C_D)$	0.71	0.81	0.76	0.80	0.83
$CI_{4,2}(C_D)$	0.81	0.62	0.61	0.74	0.69
$CI_{1,2}(L_c)$	0.71	0.74	0.73	0.72	0.74
$CI_{2,3}(L_c)$	0.86	0.84	0.87	0.74	0.82
$CI_{3,4}(L_c)$	0.83	0.88	0.80	0.82	0.82
$CI_{1,2}(P_q)$	0.84	0.78	0.82	0.87	0.85
$CI_{2,3}(P_q)$	0.79	0.75	0.78	0.86	0.88
$CI_{3,4}(P_q)$	0.77	0.84	0.81	0.86	0.91
$CI(n)$	1.07	1.08	1.100	1.100	1.08
CI	0.42	0.38	0.37	0.48	0.51
PI/s	229740	212245	236980	196572	184634
Improvement ratio	26.3%	22.5%	20.7%	33.8%	34.4%

TABLE 8: Normalized data for CI and the improvement ratio.

Normalized CI	0.357	0.071	0.000	0.786	1.000
Normalized improvement ratio	0.409	0.131	0.000	0.956	1.000

typical, the conclusions may still not be transferable to other intersections due to the complexity and diversity of real-life situations. Future research is necessary to develop statistical models to quantitatively assess the generality of these findings.

TABLE 9: Statistical results of F-test and t-test.

Sample size	Significance level	F-test		t-test	
		F_c	f	t_c	t
5	0.95	0.16	0.90	2.31	−0.20

F_c is the one-tailed critical value for the F-test; t_c is the two-tailed critical value for the t-test.

Acknowledgments

References

[1] L. T. Lin and S. M. Tsao, "A system approach on signal grouping for areawide control of computerized traffic system," in *Proceedings of the 79th Annual Meeting of the Transportation Research Board*, pp. 1–21, Transportation Research Board, Washington, DC, USA, 2000.

[2] C. Ding, W. H. Wang, X. Wang et al., "A neural network model for driver's lane-changing trajectory prediction in urban traffic flow," *Mathematical Problems in Engineering*, vol. 2013, Article ID 967358, 8 pages, 2013.

[3] D. F. Ma, D. H. Wang, Y. M. Bie et al., "A method of signal timing optimization for spillover dissipation in urban street networks," *Mathematical Problems in Engineering*, vol. 2013, Article ID 580546, 9 pages, 2013.

[4] R. J. Walinchus, "Real-time network decomposition and subnetwork interfacing," *Highway Research Record*, vol. 366, pp. 20–28, 1971.

[5] Z. Y. Liu, Q. Meng, and S. A. Wang, "Speed-based toll design for Cordon-based congestion pricing scheme," *Transportation Research C*, vol. 31, pp. 83–89, 2013.

[6] Z. Y. Liu, Y. D. Yan, X. B. Qu et al., "Bus stop-skipping scheme with random travel time," *Transportation Research C*, vol. 35, pp. 46–56, 2013.

[7] K. Lu, *Research on the foundational theory and key techniques of coordinate signal control in urban traffic network [Ph.D. thesis]*, South China University of Technology, Guangzhou, China, 2010.

[8] J. Y. Park, *Network-wide signal control with distributed Real-time travel data [Ph.D. thesis]*, University of California at Irvine, Irvine, Calif, USA, 2009.

[9] L. L. Zang, X. X. Meng, L. Jia, and L. Zhao, "An optimal control framework of traffic signals based on dynamical subarea partition for urban main-roads," in *Proceedings of the 8th World Congress on Intelligent Control and Automation (WCICA '10)*, pp. 4533–4538, IEEE, Jinan, China, July 2010.

[10] Q. Meng and Z. Liu, "Impact analysis of cordon-based congestion pricing on mode-split for a bimodal transportation network," *Transportation Research C*, vol. 21, no. 1, pp. 134–147, 2012.

[11] H. Etemadnia, K. Abdelghanya, and A. Hassana, "Network partitioning methodology for distributed traffic management applications," *Transportmetrica A*, 2013.

[12] H. N. Yagoda, E. H. Principe, C. E. Vick, and B. G. Leonard, "Subdivision of signal systems into control areas," *Traffic Engineering*, vol. 43, no. 12, pp. 42–46, 1973.

[13] E. C. P. Chang, "Evaluation of interconnected arterial traffic signals," *Transportation Planning Journal Quarterly*, vol. 1, no. 15, pp. 137–156, 1986.

[14] J. E. Moore and P. P. Jovanis, "Statistical designation of traffic control subareas," *Journal of Transportation Engineering*, vol. 111, no. 3, pp. 208–223, 1985.

[15] L. T. Lin and H. J. Huang, "A linear model for determining coordination of two adjacent signalized intersections," *Journal of Modeling in Management*, vol. 4, no. 2, pp. 162–173, 2009.

[16] K. Lu, J. Xu, and Y. Li, "Division method of coordinated control subareas based on correlation degree analysis," *Journal of South China University of Technology*, vol. 37, no. 7, pp. 6–9, 2009 (Chinese).

[17] Y. Bie, D. Wang, Q. Wei, and D. Ma, "Development of correlation degree model between adjacent signal intersections for subarea partition," in *Proceedings of the 11th International Conference of Chinese Transportation Professionals (ICCTP '11)*, pp. 1170–1180, American Society of Civil Engineers, Nanjing, China, August 2011.

[18] D. I. Robertson, "TRANSYT A traffic network study tool-Report," Tech. Rep. 253, TRRL Laboratory, London, UK, 1969.

[19] Transportation Research Board, *Highway Capacity Manual*, National Research Council, Washington, DC, USA, 2000.