



Traffic information interpolation method based on traffic flow emergence using swarm intelligence

Satoshi Suga¹ · Ryu Fujimori¹ · Yuji Yamada¹ · Fumito Ihara¹ · Daiki Takamura¹ · Ken Hayashi¹ · Satoshi Kurihara²

Received: 13 May 2022 / Accepted: 8 December 2022 / Published online: 3 January 2023
© The Author(s) 2023

Abstract

Traffic congestion has become one of the most pressing social problems in today's society, and research into appropriate traffic signal control is actively underway. At present, most traffic signal control methods define traffic signal parameters on the basis of traffic information such as the number of passing vehicles. Installing sensors at a vast number of intersections is necessary for more precise and real-time adaptive control, but this is unrealistic from the viewpoint of cost. As an alternative, we propose a swarm intelligence-based methodology that creates routes with a similar traffic volume using the traffic information from intersections already equipped with sensors and interpolates this information in the intersections without sensors in real time. Our simulation results show that the proposed methodology can effectively create similar traffic routes for main traffic flows with high traffic volumes. The results also show that it has an excellent interpolation performance for heavy traffic flows and can adapt and interpolate to situations where traffic flow changes suddenly. Moreover, the interpolation results are highly accurate at a road link where traffic flows confluence. We also developed an interpolation algorithm that is adaptable to traffic patterns with confluence traffic flows. Experiments were conducted with a simulation of merging traffic flows and the proposed method showed good results.

Keywords Swarm intelligence · Ant colony optimization (ACO) · Intelligent transport systems (ITS) · Multi-agent systems

1 Introduction

Traffic congestion is one of the most pressing social problems in society today because it can lead to environmental pollution, time loss to drivers, and economic loss. As one of the causes of traffic congestion is traffic light control that cannot adapt to dynamic traffic flow changes, one way to reduce traffic congestion is to properly control traffic

signal parameters. For this reason, there has been extensive research on traffic signal control to reduce traffic congestion in recent years.

In Japan today, traffic signals are controlled by two major methods. One is point control. This method operates by repeating patterns of control parameters prepared in advance. With this method, it is difficult to respond appropriately when unanticipated traffic flow patterns occur due to accidents or disasters.

The other is centralized control. This method utilizes parameter patterns, as with point control, but multiple patterns are designed instead of just one, and in operation, the control parameters are selected by the traffic control center on the basis of traffic information obtained from vehicle-sensing sensors installed on the road. In Japan, a method called MODERATO [1] is applied to traffic light control only at major intersections. Several centralized control methods exist internationally, including SCOOT [2], SCAT [3], and OPAC [4]. In all these methods, parameters are designed in advance, and traffic information is collected at the control center using vehicle-sensing sensors and probe information to estimate parameters in batches.

This work was presented in part at the joint symposium of the 27th International Symposium on Artificial Life and Robotics, the 7th International Symposium on BioComplexity, and the 5th International Symposium on Swarm Behavior and Bio-Inspired Robotics (Online, January 25-27, 2022).

✉ Satoshi Suga
ssuga@keio.jp
Satoshi Kurihara
satoshi@keio.jp

¹ Graduate School of Science and Technology, Keio University, Kanagawa, Japan

² Faculty of Science and Technology, Keio University, Kanagawa, Japan

The centralized control method determines parameters on the basis of traffic information and thus adapts control to the current traffic conditions. However, because the control is centralized at the control center, the computation amount increases when the control area is broad or at intersections with many crossroads, and it takes time for the control system to reflect the results. This makes the system less responsive to dynamically changing traffic conditions at individual intersections.

For this reason, autonomous distributed control systems have been attracting interest as a method for traffic signal control in recent years. In this approach, a computer unit is placed at each intersection, and each unit autonomously determines the traffic signal control parameters using the traffic information around that intersection.

Examples of autonomous distributed control include methods using mixed integer linear programming problems [5] or evolutionary strategies [6]. However, these have only been validated on fixed control areas or small road networks.

There are also methods that utilize deep reinforcement learning to learn the appropriate signal switching method for the traffic signal units at each intersection [7–10]. These methods have demonstrated excellent performance, but since they have the potential to perform fast display switching control, it would be necessary to set restrictions on the control for safety in real-world operation [11].

Another study proposed a highly responsive autonomous distributed traffic signal control system with low computational and communication costs for real-world implementations [12]. To apply such a method in a real environment, however, vehicle-sensing sensors need to be installed at intersections to obtain traffic information. As these sensors are required at a vast number of intersections to ensure accurate and real-time adaptive traffic signal control, the cost of installing them in a real environment becomes prohibitive. We, therefore, need a way to interpolate traffic information at intersections where vehicle-sensing sensors cannot be installed. In addition, real-time interpolation is necessary to take advantage of autonomous distributed traffic signal control systems that aim for immediacy.

2 Traffic volume prediction

One of the techniques for interpolating unknown data is prediction technology. Many methods for traffic volume prediction that utilize machine learning models have been proposed. For example, focusing on the fact that peak and off-peak periods in traffic conditions are time-series data that show a strong seasonal pattern, a traffic volume prediction method using a seasonal ARIMA model was proposed [13]. There are also prediction models that apply nonparametric methods, such as using a regression model [14] or a support vector machine

model [15]. Recent years have seen an increase in traffic volume prediction models based on deep learning approaches [16]. Zhao et al. [17] proposed a traffic volume prediction model using an LSTM model that outperformed the ARIMA, SVM, and RNN models.

In general, when environmental changes occur due to accidents or road construction, the traffic flow also changes. Such changes in traffic flow are abrupt and may involve unknown patterns. Therefore, traffic volume prediction based on machine learning faces the challenge of how to adapt to sudden changes in traffic flow. Moreover, since the machine learning-based approaches use the past traffic volume data at the prediction target location to predict the future traffic volume, it is necessary to acquire traffic volume data by installing vehicle-sensing sensors at the prediction target location. In locations where vehicle-sensing sensors cannot be installed, it is difficult to obtain past traffic volume data in real time, which makes it difficult to apply the time-series prediction model.

Studies have also been conducted to interpolate traffic volumes at unobserved points using information obtained from sources other than the road links to be interpolated. For example, there is a method to interpolate traffic information at unobserved points using geographic information system (GIS) datasets and the geostatistical kriging method [18]. Another method utilizes an advanced kriging approach that attempts to spatially interpolate traffic volumes at unobserved locations using regression based on Euclidean distance, network distance, the number of lanes, speed limits, and other site attributes [19]. Real-time interpolation is difficult for these interpolation methods because the objective variable is average annual daily traffic (AADT), and interpolation is performed from several years of data.

Methods for real-time interpolation have also been proposed. Aljamal et al. [20] developed a method for real-time interpolation of traffic volumes around intersections by utilizing not vehicle-sensing sensors but rather probe data obtained using Kalman filtering techniques. Probe data are obtained from vehicles equipped with dedicated probe terminals that provide traffic information such as travel speed and location. For accurate interpolation, many vehicles need to be equipped with probe terminals, but only a low percentage of vehicles in Japan are thus equipped. Therefore, in this study, we propose a real-time and dynamic traffic information interpolation system for roads without vehicle-sensing sensors using only information that can be obtained from roads with other vehicle-sensing sensors.

3 Traffic information interpolation system based on swarm intelligence

3.1 System overview

As stated earlier, it is not possible to obtain real-time traffic information at intersections where vehicle-sensing sensors are not installed. Therefore, our proposed system interpolates the traffic information of road links without vehicle sensors in real-time based on the traffic information of each road link where vehicle sensors are installed. Specifically, it creates a route between road links with similar traffic volume from the traffic information of each road link that can be obtained and then interpolates the traffic information of road links without vehicle-sensing sensors based on the newly created route. The traffic information includes the traffic volume per unit time for each road link. We use the ant colony optimization (ACO) algorithm [21] to create routes between road links with similar traffic volumes.

As the ACO algorithm is robust and can adapt to dynamic changes in the environment, its flexible routing capabilities have been applied to a variety of tasks, such as solving delivery route optimization and dynamic job-shop scheduling problems [22, 23]. There is also a method using the ACO algorithm to implement path planning for autonomous mobile robots [24]. The ACO algorithm has demonstrated excellent results as a path optimization method with dynamic changes and is considered effective as a path finding method for traffic flow with changes.

3.2 Agent behavior network

Our objective is to create a route featuring similar traffic volumes based on the ACO algorithm. Therefore, we need to prepare a network in which the ant agents can act. The network we constructed is based on the road environment, where nodes represent road links and edges represent connections between road links (Fig. 1). The pheromone trail changes in accordance with the similarity of the traffic flow. The ant agents carry traffic information on such a network to create routes between similar traffic volumes.

3.3 Traffic information data flow

In traffic signal control by autonomous distributed control, the traffic volume information per cycle length of signal control is used for computation. Therefore, the proposed method interpolates the traffic volume information once per cycle. The traffic information on the road link where a sensor is installed is sent to the traffic information interpolation system at the end of one signal control cycle. At the same

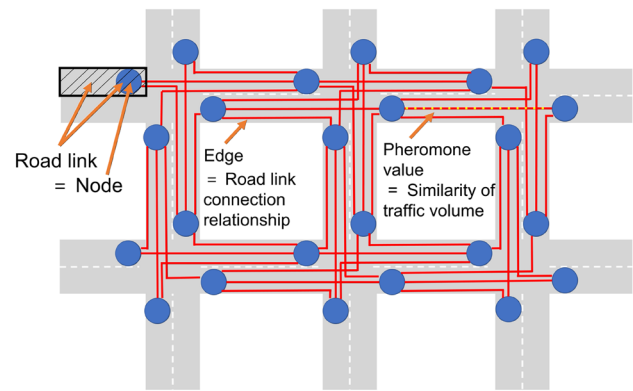


Fig. 1 Structure of pheromone map

time, to improve adaptability to changes in traffic flow, the pheromone trail updates by the ACO algorithm (defined in 3.2) are performed repeatedly during one cycle of signal control. The pheromone value in the pheromone trail is not initialized when the signal control advances to the next cycle step, and the pheromone trail state continues.

3.4 Interpolation system flow

The following subsections describe the details of the traffic information interpolation system using the ACO algorithm. First, ant agents are generated on a road link where traffic information is available. These agents have the traffic information of that road link. Next, each agent decides the movement route based on the amount of pheromone on the pheromone trail and propagates the traffic information it possesses to the destination road link. The agent moves by choosing a route based on the pheromone trail until it reaches a road link where more traffic information is available. After the movement, the agent compares its departure traffic information with that of the current road link and evaluates the movement route. Then, on the basis of the evaluation value, the pheromone value on the moving route is updated. We attempted to create similar traffic flow routes and interpolate appropriate traffic information by repeating this process. Since the traffic flow changes over time, the flexibility and adaptability of the ACO algorithm are effective properties. Figure 2 shows the overall flow of the traffic information interpolation system.

3.4.1 Generate agents

As stated earlier, ant agents are initially generated on each road link where traffic information is available. We define a signal cycle as c , a step of the ACO algorithm as t , a road link as i , and the number of agents generated at step t as $N_i(t)$. The moving average of the traffic volume information

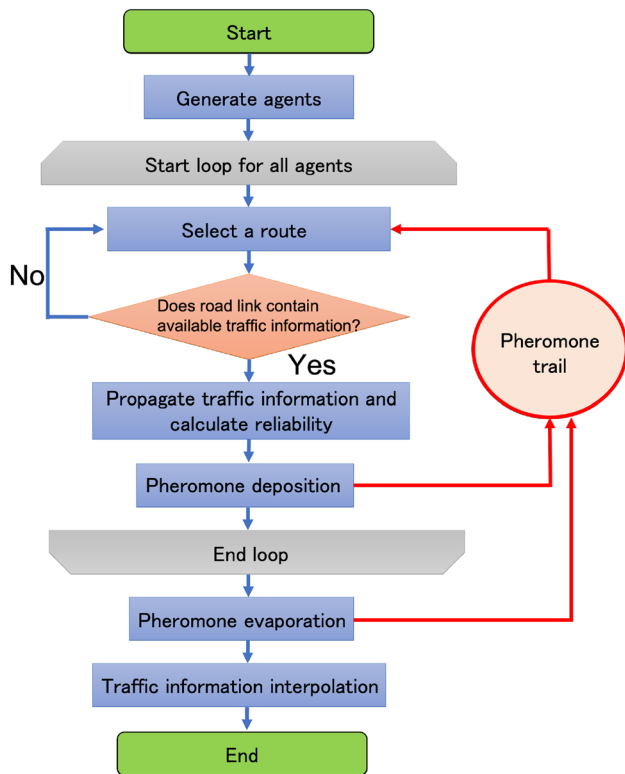


Fig. 2 Overview of the traffic information interpolation system

of a road link in a signal cycle c is defined as $\overline{RV}_i(c)$, which in this study is the moving average of the last five cycles of traffic volume at road link i . A moving average is calculated and smoothed to reduce the range of prediction errors per cycle. The agent has the role of propagating this information to other road links. The number of agents $N_i(t)$ to be generated is determined on the basis of the moving average $\overline{RV}_i(c)$ of the traffic information of the road link:

$$N_i(t) = \overline{RV}_i(c) \tag{1}$$

The reason for applying this as the number of ant agents that generate traffic volume information is to make interpolation concerning heavy traffic flows more accurate. When there are more ant agents, the amount of pheromone increases, and this combined with the increase in the amount of pheromone due to traffic similarity results in a stronger emergence of paths between similar traffic volumes in main traffic flows. In this study, Eq. (1) is applied to facilitate the emergence of similar traffic flow paths at locations with high traffic volumes.

3.4.2 Select a route

The next step is to determine the route of each agent. Each agent’s destination road link choice is a road link that it can proceed to from its current location. For example, if the agent is on a road link that connects to a four-way intersection, three road link options are listed: go straight ahead, turn right, and turn left. The edges on which the agent proceeds have a pheromone value indicating the similarity of traffic information between the start road and the connection road links. This pheromone value is increased by pheromone addition performed by the agents. The evaporation of the pheromone reduces this value. We set the initial pheromone value at system start time 0.1 for all edges. Each agent prioritizes the route with the high pheromone value. In the case of road link i to destination road link j , if we define the pheromone value between i and j as $\tau_{i,j}(t)$, the movement probability of the agent from road link i to j is defined as

$$p_{i,j}(t) = \frac{\tau_{i,j}(t)}{\sum_{j \in J} \tau_{i,j}(t)}, \tag{2}$$

where J denotes the set of road links that can be moved to from road link i . The agent randomly selects one road link from among the candidates for the destination road link with a certain probability, regardless of the route selection by Eq. (2). This selection is designed to prevent excessive convergence of the pheromone trail, which would prevent the discovery of routes to other road links with high similarity. The agent follows this method to perform route selection and movement until it reaches another road link where traffic information is available.

3.4.3 Propagate traffic information and calculate reliability

The traffic information interpolation is based on the reliability of the traffic information propagated by the agents (details are explained in 3.4.6). Here, we explain how to calculate the reliability of the propagation data. After determining the route, the agent moves to the destination road link and propagates its traffic information to the road link where traffic information cannot be obtained. At the same time, it calculates the reliability of its traffic information. In general, the traffic information possessed by agents generated on road links close to the road link to be interpolated is reliable as interpolated values. In addition, the traffic information between road links with high pheromone value edges is considered to have high reliability thanks to being similar. Therefore, the factors involved in the reliability level are the distance moved and the pheromone value of the moving route.

When agent k , which started from road link i , is passing through road link m , and the distance it has moved is $h_{i,m}$, the interpolated traffic information $AV_m^k(c)$ and the reliability level $AR_m^k(c)$ it possesses are calculated as

$$AV_m^k(c) = \overline{RV_i(c)}, \tag{3}$$

$$AR_m^k(c) = (1 - Hreduce)^{h_{i,m}} \times \frac{\sum_{n=1}^{h_{i,m}} \tau_{p_{n-1},p_n}(t)}{h_{i,m}}, \tag{4}$$

where $Hreduce$ is a parameter indicating the rate of decrease in the reliability of the agent’s travel movement, which is $Hreduce \in (0, 1]$. In this study, we set $Hreduce = 0.05$. In Eq. (4), $\sum_{n=1}^{h_{i,m}} \tau_{p_{n-1},p_n}(t)$ represents the sum of the pheromone values of each edge that the agent has passed from the starting road link to road link m when it passes road link m at step t .

3.4.4 Pheromone deposition

The agent follows the route it has traveled while depositing the pheromone and returns to the road link from which it was generated. The amount of pheromone deposited is determined by the similarity between the traffic information at the agent’s starting road link and that at the destination road link. When the agent generated on road link i moves to road link p , the pheromone increase $\Delta\tau_{i,p}(t)$ in the route traveled by this agent is calculated as

$$d = |\overline{RV_i(c)} - \overline{RV_p(c)}|, \tag{5}$$

$$\Delta\tau_{i,p}(t) = \frac{1}{1 + d}, \tag{6}$$

where d is the Euclidean distance between the traffic information of the two road links. When traffic information is similar between roads, more pheromones are deposited.

3.4.5 Pheromone evaporation

Pheromones evaporate at the rate of $e(e \in (0, 1])$ in one creation phase. This evaporation effect lowers the pheromone value between road links with low similarity, and traffic information can be propagated efficiently between road links with high similarity. In this study, we set $e = 0.05$:

$$\tau_{i,j}(t + 1) = \tau_{i,j}(t) \times (1 - e) \tag{7}$$

3.4.6 Traffic information interpolation

When all agents have finished moving, each road link adopts the most reliable interpolated traffic information among the

candidates propagated to itself as the interpolated traffic information. In road link m , the number of propagated interpolated traffic information candidates is o , and their information is $AV_m^1(c), AV_m^2(c), \dots, AV_m^o(c)$ and $AR_m^1(c), AR_m^2(c), \dots, AR_m^o(c)$. The interpolated value $PV_m(c)$ is calculated as

$$PV_m(c) = AV_m^l(c) (l = \text{argmax} AR_m^k(c)) (k = 1, 2, \dots, o). \tag{8}$$

Through the above procedure, the traffic information of road links without sensors is interpolated.

3.5 Interpolation method for road link where traffic flows confluence

The method proposed in the previous sections considers the difference in traffic volume between two road links as a pheromone value to form traffic flow routes with similar traffic volumes. If two or more traffic flows merge at an intersection and the traffic volume at the merging road links needs to be interpolated, the difference in traffic volume between the two road links will be the difference between the traffic volume before and after the merge, and the pheromone trail will not form well. For example, in Fig. 3a, the main traffic flows are independent. Therefore, the traffic volume on road links A and B is similar, and that on road links C and D is similar, thus forming a pheromone trail that can be applied by the method proposed in the previous sections. In contrast, the main traffic flows merge in Fig. 3b, so the traffic volume on road link G is the combined volume of traffic on road links E and F. Because of the large difference in traffic volume between road link E (F) and road link G, the pheromone trail cannot be formed, even though the main traffic flow originally continues from road link E (F) to G. For this reason, it is difficult to detect the main traffic flow after the merge, making it challenging to interpolate the points marked with a star in the figure. Therefore, in this section, we present a traffic information interpolation method that can be adapted to cases where traffic flows converge.

3.5.1 Route memorization

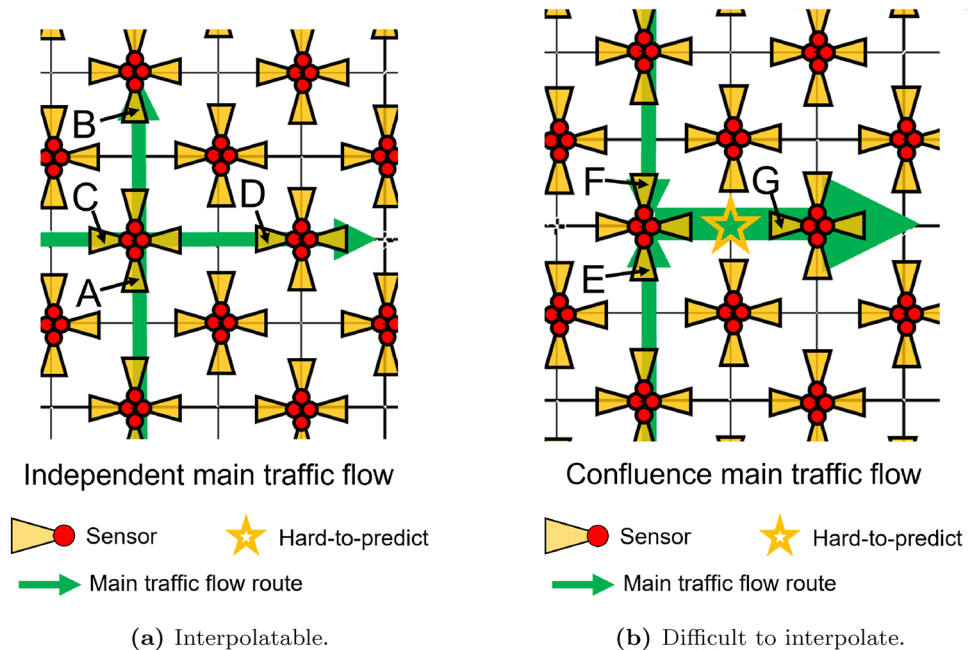
First, each agent memorizes the routes it has traversed. If p_n is the road link that the agent has traveled to, the agent has the following information:

$$Path = \{p_0, p_1, \dots, p_n\}. \tag{9}$$

3.5.2 Pheromone deposition considering confluence of traffic flow

The similarity calculation in 3.4.4 is changed to a formula that takes into account the confluence of traffic flow.

Fig. 3 Two traffic flow patterns



If an agent has information on a transit route *Path* and moves to road link p_n , the set of all agents that have moved to road link p_n is $Agent_{p_n}$, and the set of agents that have passed along the same route as the agent is $Agent_{Path}$. Then, the pheromone increase $\Delta\tau_{i,p}(t)$ in the route traveled by this agent is calculated as

$$d = \begin{cases} |\overline{RV_i(c)} - \overline{RV_{p_n}(c)}| & (\overline{RV_i(c)} \geq \overline{RV_{p_n}(c)}) \\ |\overline{RV_i(c)} - (\overline{RV_{p_n}(c)} \times \frac{n(Agent_{Path})}{n(Agent_{p_n})})| & (\overline{RV_i(c)} < \overline{RV_{p_n}(c)}) \end{cases}$$

$$\Delta\tau_{i,p}(t) = \frac{1}{1 + d}, \tag{10}$$

where $n(Agent_{p_n})$ indicates the number of agents that have moved to the road link p_n and $n(Agent_{Path})$ indicates the number of agents that have passed through the same path as agent k . From Eq. (10), the number of vehicles only on the agent’s route is estimated by multiplying the traffic volume on the destination road link by the percentage of agents that have passed along the same route.

3.5.3 Traffic information propagation and prediction calculation considering confluence of traffic flow

By considering the number of agents with respect to the route, we attempt to calculate an interpolation value that is close to the post-merge traffic volume. To obtain the necessary information for this purpose, when agents propagate traffic information, they propagate the information $AP_m^k(c)$, which is the agent’s travel route up to road link m , in addition to the interpolated traffic information $AV_m^k(c)$ and confidence level $AR_m^k(c)$ (discussed in 3.4.3), as

$$AP_m^k(c) = \{p_0, p_1, \dots, p_m\}, \tag{11}$$

In the calculation of the traffic interpolation values, o candidate interpolation values are propagated to the road link m (discussed in 3.4.6). Their information is $AV_m^1(c) \dots AV_m^o(c)$, $AR_m^1(c) \dots AR_m^o(c)$ and $AP_m^1(c) \dots AP_m^o(c)$. The interpolated value $PV_m(c)$ is calculated as

$$PV_m(c) = AV_m^l(c) \times \frac{o}{n(AP_m^l(c))} \tag{12}$$

$$(l = \text{argmax} AR_m^k(c) (k = 1, 2, \dots, o),$$

where $n(AP_m^l(c))$ indicates the number of propagation information that is the same path as the l -th transit route information. The total number of propagation information is divided by this value and multiplied by the estimated traffic volume to estimate the volume of traffic after confluence.

3.5.4 Increase in the number of agents generated

The extended interpolation algorithm that takes into account the merging of main traffic flows calculates pheromone deposit values and interpolation values based on the ratio of ant agents that followed the same path, as shown in Eqs. (10) and (12). In this case, although the ant agent takes into account the pheromone value, it chooses its route by probability. Therefore, the number of agents must be increased to eliminate fluctuations in path selection due to probability and to perform proper pheromone deposition. The greater the number of agents, the more appropriate the agent path selection according to the ratio of pheromone values. We, therefore, increase the number of

agents by multiplying the number of agents generated by n times the moving average of traffic information on road links:

$$N_i(t) = \overline{RV_i(c)} \times n. \tag{13}$$

In this study, we set $n = 5$.

4 Evaluation experiment

4.1 Performance evaluation

We evaluate the effectiveness of the proposed system with experiments using the Simulation of Urban Mobility (SUMO) [25] traffic simulator. Specifically, we examine the accuracy of traffic information interpolation for a road link that flows into an intersection with no sensors, which is the main traffic flow with high traffic volume. The root mean square error (RMSE) is calculated for the actual value obtained by the simulator. The range of the evaluation target is set to within the range of the agent’s behavior, as shown in Fig. 4.

4.2 Comparison methods

We assume the algorithms discussed here will be applied in the real world, so we used the following three methods, which are considered applicable in the real world, as comparison methods.

4.2.1 Traffic volume survey

In Japan, a traffic survey is conducted once every five years to obtain the hourly traffic volume of a road link on a particular day. Traffic volume surveys are generally conducted under normal traffic flow conditions, not under special situations such as accidents. Since this experiment uses virtual simulation, we conducted simulations under the same conditions as normal traffic flow to simulate a traffic volume survey, and 1-h traffic volume data for all road links were obtained. Under the same situation as normal traffic flow means that the parameters of the vehicle routes that form the main traffic flow, the vehicle inflow probability, and the vehicle route choice probability (described in detail in 5.1) are the same as in the simulations conducted in 5.1. The interpolated value in the comparison method is the average traffic volume for one cycle, calculated by dividing the hourly traffic volume for the target time of the road link to be interpolated by the number of cycles per hour.

4.2.2 Clustering

From the traffic survey data values for each road link obtained in 4.2.1, k-means++ [26] is used to group road links with similar traffic flow patterns. In the interpolation scene, based on the result of clustering, the average value of real-time traffic volume data acquired from sensor-installed road links that belong to the same cluster as the road link to be interpolated is applied as the interpolated value.

4.2.3 Neighbor interpolation

Traffic volume surveys are based on historical statistics. The clustering method uses real-time traffic as interpolated values, but the clustering calculation itself relies on historical statistical data. Therefore, we use the neighbor interpolation method as an interpolation method utilizing real-time traffic volume data. The experiment evaluates the results of traffic information interpolation for road links on main traffic flows. Therefore, in the neighbor interpolation method, the average value per five cycles on the road link with the largest traffic volume among adjacent road links within two hops of the road link to be interpolated is applied as the interpolated value. The reason for taking a moving average of five cycles is to reduce the range of prediction error per cycle, as with the proposed method.

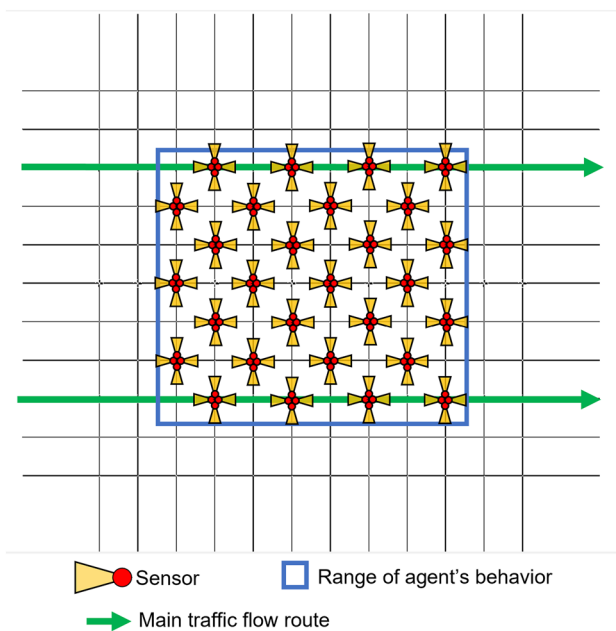


Fig. 4 Simulation of normal traffic flow

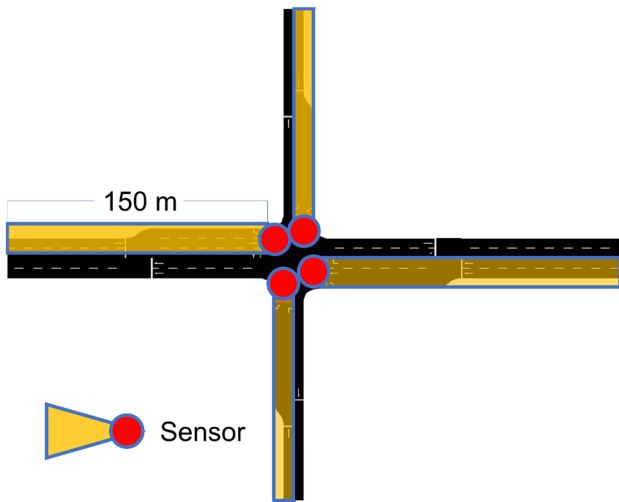


Fig. 5 Sensor installation details

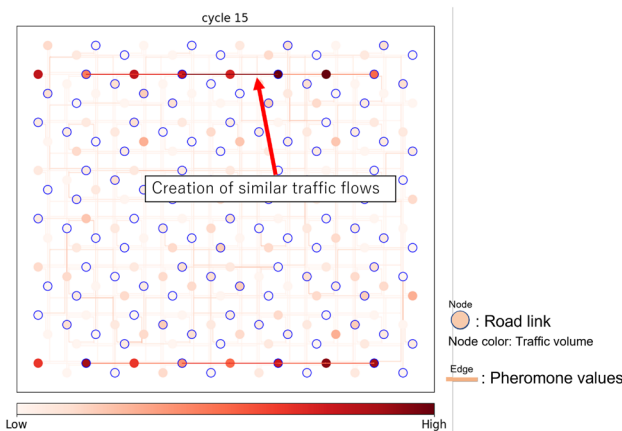


Fig. 6 Creation of routes between similar traffic volumes

5 Results

5.1 Experiment in normal traffic flow

First, we performed interpolation experiments in normal traffic flow. The vehicles that form the main traffic flow travel along the defined route shown in Fig. 4. During the entire simulation period, the main traffic flow routes remain unchanged. Therefore, the road links for accuracy evaluation in this experiment are those on the main traffic flow (indicated by the green arrows) in Fig. 4, where no sensors are installed. The layout of the road links where the sensors are installed is shown in Fig. 4. The sensors are assumed to be installed near traffic lights and acquire traffic volumes for each road link entering the intersection. The acquisition range of the sensors is

Table 1 Average RMSE of interpolation values in normal traffic flow

Method	Avg. RMSE
Traffic volume survey	3.91
Clustering	3.39
Neighbor interpolation	2.84
Proposed	2.88

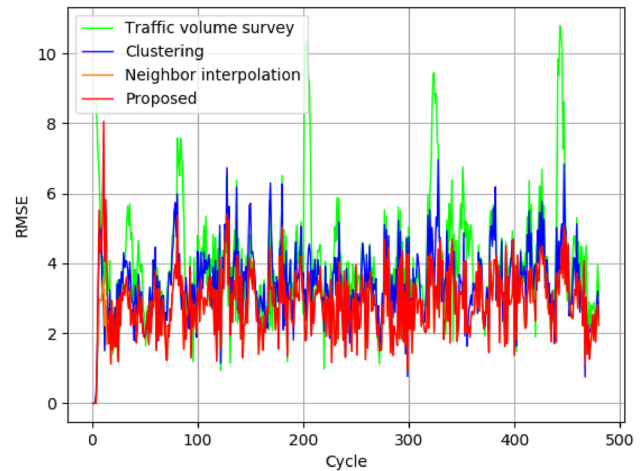


Fig. 7 Interpolation results in normal traffic flow

assumed to be 150 m. Figure 5 shows an enlarged image of one intersection from Fig. 4.

We set the congestion time for each simulation period. The ratio of vehicle inflow during congestion and non-congestion times is set to 2 : 1. Vehicles form the main traffic flow, but some vehicles also choose other routes with probability at the intersection. The inflow ratio between these vehicles and the vehicles that form the main traffic flow is set to 1 : 6.

5.2 Creation of routes between similar traffic volumes

The results for similar traffic volumes created by the proposed method are shown in Fig. 6. Nodes indicate road links, and edges are placed between road links that can be proceeded on. The color of the node indicates the traffic volume in one cycle, where a darker color means a greater amount of traffic. The color of the edge indicates the pheromone value, where the higher the pheromone value, the darker the color. As we can see in the figure, routes with high pheromone values were formed along the nodes with high traffic volume. This demonstrates that the proposed method can create routes that connect similar traffic volumes.

However, in Fig. 6, no similar paths could be generated for the left-most node and its right neighbor among the nodes that represent the main traffic flow. This is because the

starting point of the road link where the sensor is installed in the main traffic flow is not the left-most road link, but the one to the right of it. In this case, the starting point of the ant agents is also not the left-most road link, but its right neighbor, and thus no similar paths could be generated between the left-most road link and the right-neighbor road link.

5.3 Interpolation results for normal traffic flow simulation

The results of traffic information interpolation are shown in Fig. 7 and Table 1. The horizontal axis in the figure shows the number of cycles, and the vertical axis shows the RMSE values. The green, blue, orange, and red lines show the interpolation results by traffic volume survey, by clustering, by neighbor interpolation, and by the proposed method, respectively. Note that the red and orange lines overlap because the neighborhood interpolation method produced similar results to the proposed method.

With the interpolation by the traffic volume survey, we can see that the RMSE values were higher when the traffic inflow changed. In contrast, since the clustering, neighbor interpolation method, and proposed method interpolate from real-time traffic data, the interpolation was adaptive to the inflow changes. In particular, since the proposed method interpolates only from the road links existing on the route with similar traffic volume, it was more accurate than the clustering method.

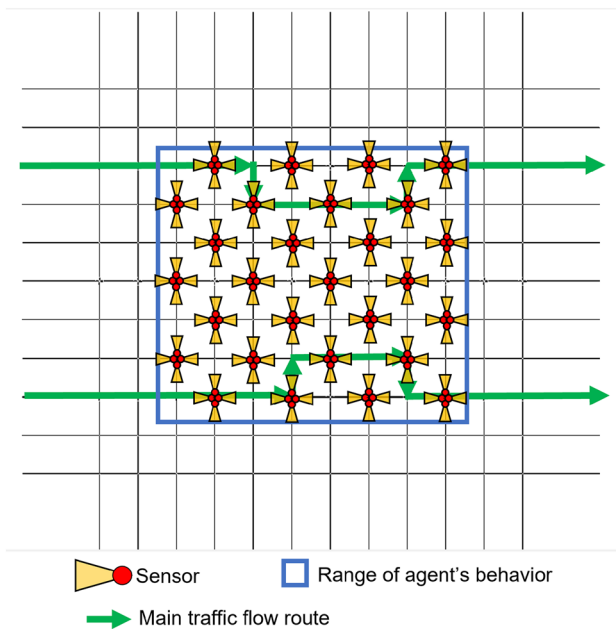


Fig. 8 Simulation of the traffic flow after sudden change

5.4 Experiment under changing traffic flow

Next, we conducted a simulation experiment to examine how well the proposed method could deal with sudden changes in traffic flow. In the first half of the simulation, we use the same main traffic flow (Fig. 4) and parameters as in 5.1. In the second half, the main traffic flow route changes from the route shown in Fig. 4 to that shown in Fig. 8. Therefore, the road links for accuracy evaluation in this experiment are those on the main traffic flow (green arrows) in Figs. 4 and 8, where no sensors are installed. The layout of the road links where the sensors are installed is shown in Fig. 8.

5.5 Results of interpolation in traffic flow with sudden change

The interpolation results in traffic flow with sudden changes are shown in Fig. 9 and Table 2. The main traffic flow varied after 240 cycles. In the first half of the simulation, there is no significant difference between the results of the different

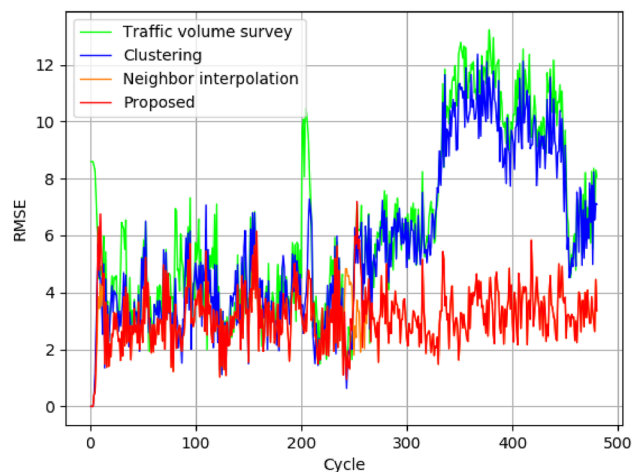


Fig. 9 Results of interpolation in traffic flow with sudden change

Table 2 Average RMSE of interpolation values in traffic flow with sudden change

Method	Traffic flow	Avg. RMSE
Traffic volume survey	Before change	4.23
	After change	8.77
Clustering	Before change	3.60
	After change	8.54
Neighbor interpolation	Before change	3.01
	After change	3.30
Proposed	Before change	3.05
	After change	3.40

methods. A normal traffic flow simulation is performed here, so the method with the traffic volume survey simulates similar parameters and even the interpolation method based on the traffic volume survey gives relatively good interpolation results. However, we can see that the RMSE values of the traffic volume survey and clustering methods increased significantly in the second half of the simulation when the traffic flow changed.

In contrast, the neighbor interpolation method and the proposed method exhibited only a slight increase in RMSE values. The traffic volume survey and clustering method were not able to adapt to sudden changes in traffic flow because there was no data for such changes. We presume the slight increase in the proposed method was due to the pheromone trail not being adapted immediately after the change. Since the neighbor interpolation method also interpolates the moving average of five steps, the RMSE is considered to have increased due to errors immediately after the traffic flow change. These interpolation results demonstrate that the proposed method can interpolate traffic information adaptively to changes in traffic flow.

5.6 Experiments under sparse sensor placement

Since the proposed method generates routes based on the traffic volume between road links where sensors are installed, it is likely to be affected by the sensor placement status. Therefore, we conducted interpolation experiments

with changing traffic flow under sparse sensor placement conditions using the sensor placement pattern shown in Fig. 10, where the sensors are not placed over a long section near the center of the road. The parameter settings for the traffic simulation are the same as in 5.4. The accuracy evaluation was performed on road links on the main traffic flow (green arrows in Fig. 10) where no sensors were installed.

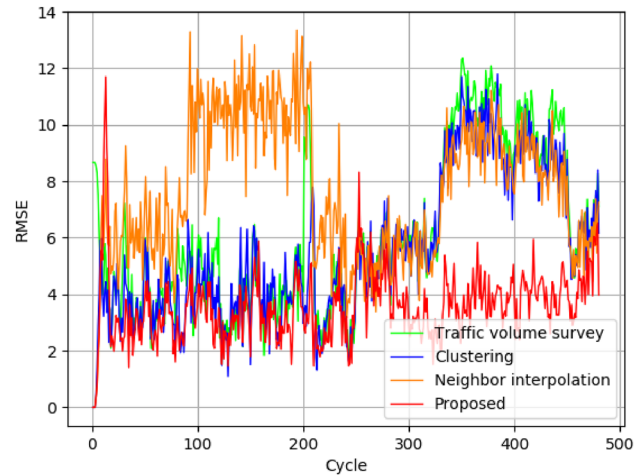


Fig. 11 Results of interpolation under sparse sensor placement

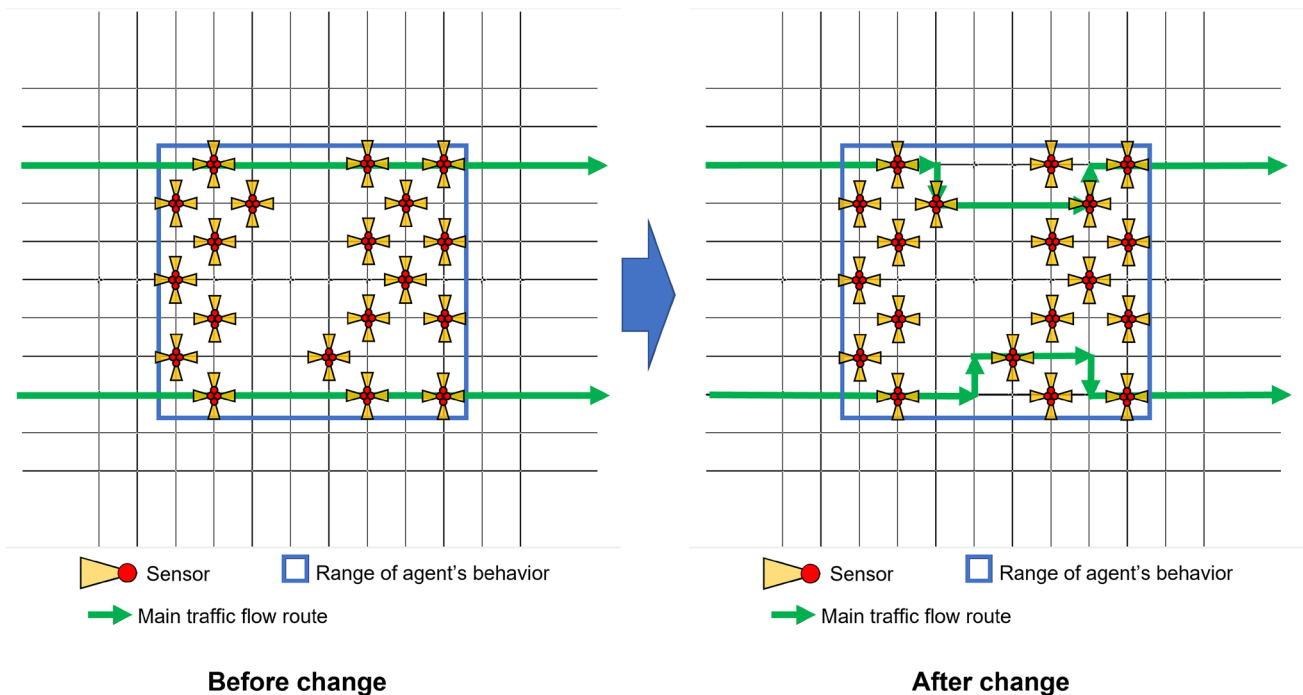


Fig. 10 Experimental setup with sparse sensor placement

Table 3 Average RMSE of interpolation values under sparse sensor placement

Method	Traffic flow	Avg. RMSE
Traffic volume survey	Before change	4.36
	After change	7.88
Clustering	Before change	3.79
	After change	7.40
Neighbor interpolation	Before change	8.18
	After change	7.16
Proposed	Before change	3.21
	After change	3.91

5.7 Experimental results for sparse sensor placement

The interpolation results under sparse sensor placement are shown in Fig. 11 and Table 3. The main traffic flow changed over 240 cycles. As in the previous experiments, the traffic volume survey and the clustering method can perform adequate interpolation before the traffic flow changes but are not able to adapt to the changes after. In the neighborhood interpolation method, the RMSE values were higher when the sensor placement was sparse, because the main traffic flow could not be detected regardless of the traffic flow change. As for the proposed method, it shows a slight increase in RMSE value under sparse sensor placement compared to the densely populated sensor condition. This can be attributed to the longer distance between the sensors, which took

longer to form an appropriate pheromone trail immediately after the change in traffic flow. Although the RMSE value increased slightly with the proposed method, it was still able to perform appropriate traffic information interpolation and adapt to changes in traffic flow.

6 Experiments under traffic flow confluence

6.1 Experimental settings

We conducted experiments to investigate the traffic information interpolation for a traffic pattern in which two traffic flows converge. The vehicles that form the main traffic flow travel along the defined route shown in Fig. 12.

There are two main traffic flows at the inflow, and at the intersection near the center, they merge to form one large traffic flow. In this simulation, the star-shaped point is a road link where traffic volumes need to be interpolated after traffic flow confluence, which is difficult to do using the interpolation method without considering confluence. As in 5.1, we set the congestion time for the simulation period. The ratio of vehicle inflow during congestion and non-congestion times is set to 2 : 1. Some vehicles also choose the route with probability at the intersection. The inflow ratio between these vehicles and the vehicles that form the main traffic flow is set to 1 : 6.

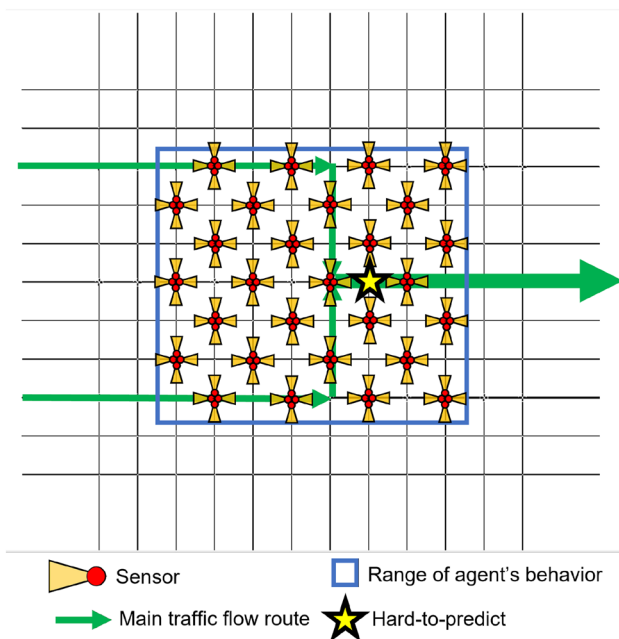


Fig. 12 Simulation of confluence traffic flow

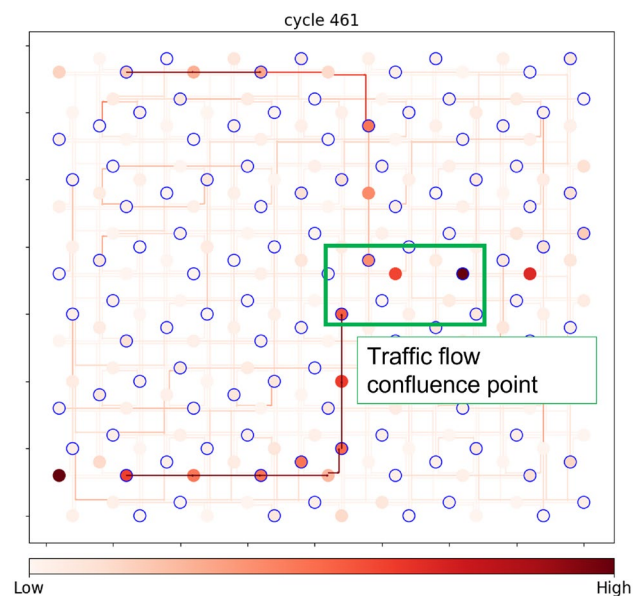


Fig. 13 Creation of route by interpolation method that does not consider confluence traffic flow

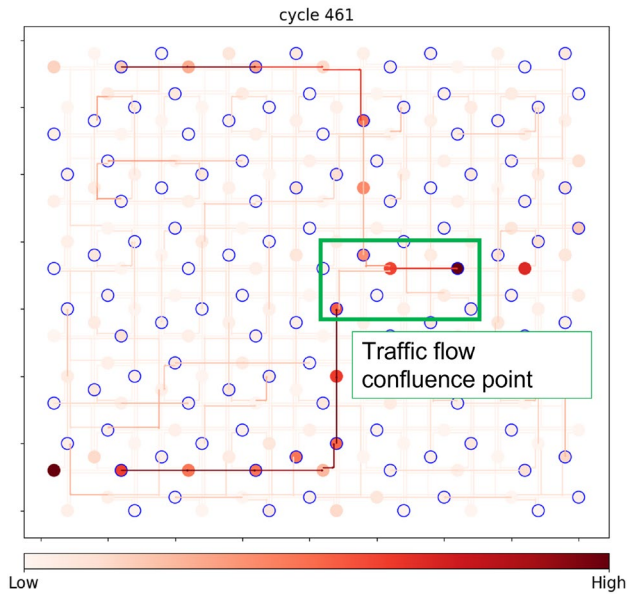


Fig. 14 Creation of route by interpolation method that does consider confluence traffic flow

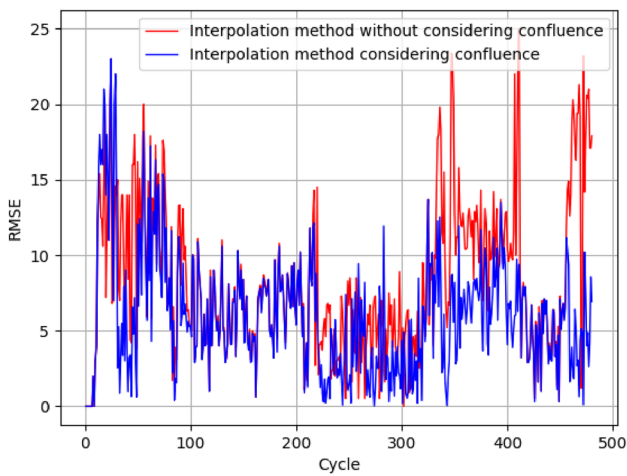


Fig. 15 Interpolation results at a road link after traffic flow confluence

Table 4 Average RMSE of interpolation values at a road link after traffic flow confluence

Method	Avg. RMSE
Method without considering confluence	9.02
Method considering confluence	6.03

6.2 Experimental results

We conducted the experiments using one interpolation method that does not consider traffic flow confluence and one that does consider it. The results for the similar traffic

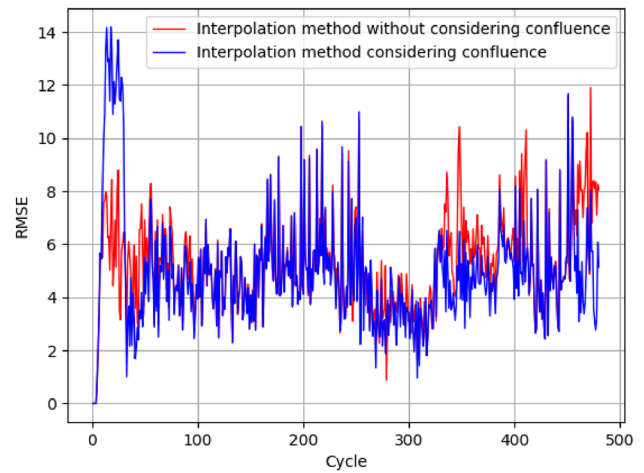


Fig. 16 Interpolation results for all interpolated road links

Table 5 Average RMSE of interpolation values for all interpolated road links

Method	Avg. RMSE
Method without considering confluence	5.40
Method considering confluence	4.96

volume route created by these two methods are shown in Figs. 13 and 14, respectively.

There is a fivefold difference in the number of agents generated by the two methods. Therefore, the maximum value of the pheromone indicated by the colors in Figs. 13 and 14 has been adjusted to 160 and 800, respectively.

As we can see in Fig. 13, the interpolation method without considering confluence can generate the main traffic flow paths up to the road links before confluence, but not after. In contrast, Fig. 14 shows that the interpolation method considering confluence can form the main traffic flow on the road link both before and after confluence. Figure 15 and Table 4 show the RMSE values for the traffic information interpolated road links after traffic flow confluence (marked with stars in Fig. 12), where the red and blue lines denote the interpolation method without and with considering confluence, respectively. As we can see, particularly in the latter half of the period when the pheromone field is stable, the interpolation method considering confluence has superior results compared to the method without. Moreover, the average RMSE values for the road links to be interpolated after traffic flow confluence (Table 4) are lower for the interpolation method considering confluence. Fig. 16 and Table 5 show the overall interpolation results, including other interpolated road links on major traffic flows with high traffic volumes. We can see here that the interpolation method considering confluence is slightly better than that without,

even when the interpolation results include road links other than those after the traffic flow confluence. Specifically, both methods showed comparable performance except for the traffic volume interpolation for the road link after the traffic flow merges, and the interpolation method considering confluence showed superior results for the traffic volume interpolation for the road link after the traffic flow merges. Thus, the average RMSE value of the interpolation method considering confluence was higher, even overall.

7 Simulation on real-world road maps

7.1 Settings

We are currently planning a real-world implementation of the proposed method with an autonomous distributed traffic signal control method. Therefore, we conducted an interpolation experiment using the simulation map shown in Fig. 17, assuming real-world operation. In this experiment, the main traffic flow changes in the first and second halves of the simulation. The black lines in the figure indicate roads and the thick lines indicate arterial roads. The green and orange arrows show the main traffic flow in the first and second halves of the simulation. The simulation period is 43,200 steps (12 h) and includes a congested time in which the vehicle inflow ratio is 2:1 (congestion time:non-congestion time). We include vehicles that travel at random in addition to the vehicles that form the main traffic flow. These vehicles flow in randomly from all entrances and travel along randomly formed routes.



Fig. 17 Simulation on real-world road maps

Table 6 Average RMSE of interpolation values in real-world road map

Method	Traffic flow	Avg. RMSE
Traffic volume survey	Before change	2.93
	After change	8.56
Clustering	Before change	3.20
	After change	7.42
Proposed	Before change	3.29
	After change	3.59

7.2 Results

We evaluated the accuracy of the traffic information interpolation for road links that form major traffic flows (green and orange arrows in Fig. 17) where no sensors were installed. RMSE is used as the evaluation metric. The interpolation results are shown in Fig. 18 and Table 6. We can see that the interpolation of the traffic volume survey was the most highly evaluated in the first half of the simulation, as the distances between intersections were not evenly spaced. This indicates that the error between the traffic volume of the road link to be interpolated and the road link on the created route was large. Since the clustering and the proposed method interpolate the values with the traffic volume of the surrounding road links, they had slightly higher RMSE values than the traffic volume survey.

In contrast, the interpolation results after traffic flow changes show that the proposed method performed better than the traffic volume survey and clustering. These two methods are not adaptable to sudden changes, while the proposed method can adaptively interpolate to sudden changes even in the simulations with the real-world road map.

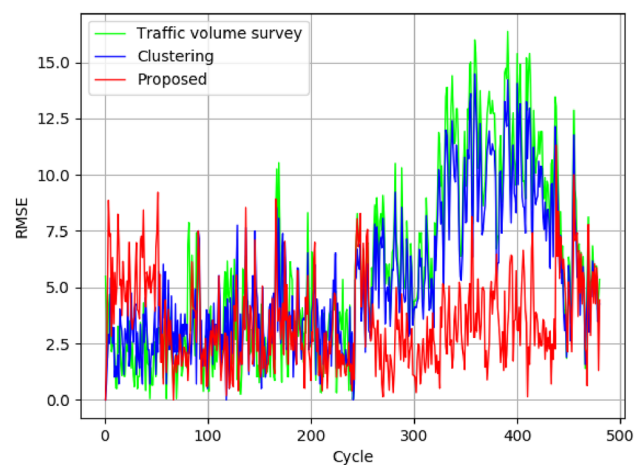


Fig. 18 Interpolation results in real-world road map

8 Conclusion

In this paper, we proposed a system for the real-time interpolation of traffic information at intersections without sensors based on a swarm intelligence algorithm. The experimental results demonstrate that the proposed method can create similar traffic routes for main traffic flows with high traffic volumes, and that it is adaptable to sudden changes in traffic flow. We also developed an interpolation algorithm that is adaptable to traffic patterns with confluence traffic flows.

Future work will include implementing a real-world application system that combines vehicle-sensing sensors, traffic signal control units, traffic signal control research, and the proposed system.

Acknowledgements This work was supported by a grant from NEDO: Realization of a Smart Society by Applying Artificial Intelligence Technologies.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Sakakibara H, Usami, T, Itakura S, and Tajima T (1999) MODERATO (Management by Origin-DEstination Related Adaptation for Traffic Optimization), proceedings 199 IEEE/IEEJ/JSAI international conference on intelligent transportation systems 38–43
2. Robertson D, Bretherton R (1991) Optimizing networks of traffic signals in real time: the SCOOT method. *IEEE Trans Veh Technol* 40(1):11–15
3. Sims A, Dobinson K (1980) The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. *IEEE Trans Veh Technol* 29(2):130–137
4. Gartner N (1990) OPAC: strategy for demand-responsive decentralized traffic signal control. *IFAC Proc* 23(2):241–244
5. Xu Y, Wu N, Li D, Xi Y (2020) Online offset optimization for urban traffic network with distributed model predictive control. *IFAC-PapersOnLine* 53(2):6604–6609
6. Zheng Y, Guo R, Ma D, Zhao Z, Li X (2020) A novel approach to coordinating green wave system with adaptation evolutionary strategy. *IEEE Access* 8:214115–214127
7. van del Pol E and Oliehoek FA (2016) Coordinated deep reinforcement learners for traffic light control, proceedings of learning, inference and control of multi-agent systems (NIPS 2016)
8. Li L, Lv Y, Wang FY (2016) Traffic signal timing via deep reinforcement learning. *IEEE/CAA J Autom Sinica* 3(3):247–254
9. Wei H, Chen C, Wu K, Zheng G, Yu Z, Gayah VV, and Li Z (2019) Deep reinforcement learning for traffic signal control along arterials, proceedings of the 2019 DRL4KDD 19
10. Oroojlooy A, Nazari M, Hajinezhad D, Silva J (2020) Universal attention-based reinforcement learning model for traffic signal control. *Adv Neural Inf Process Syst* 33:4079–4090
11. Chu T, Wang J, Codecà L, Li Z (2019) Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Trans Intell Transp Syst* 21(3):1086–1095
12. Kurihara S, Ogawa R, Shinoda K, Suwa H (2016) Proposed traffic light control mechanism based on multi-agent coordination. *J Adv Comput Intell Inform* 20(5):803–812
13. Kumar SV, Vanajakshi L (2015) Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *Eur Transp Res Rev* 7(3):1–9
14. Brian LS, Michael JD (1997) Traffic flow forecasting: comparison of modeling approaches. *J Transp Eng* 123(4):261–266
15. Zhang Y, Liu Y (2009) Traffic forecasting using least squares support vector machines. *Transportmetrica A* 5(3):193–213
16. Lingras P, Sharma S, Zhong M (2002) Prediction of recreational travel using genetically designed regression and time-delay neural network models. *Transp Res Rec* 1805:16–24
17. Zhao Z, Chen W, Wu X, Chen PCY, Liu J (2017) LSTM network: a deep learning approach for short-term traffic forecast. *IET Intell Transp Syst* 11(2):68–75
18. Selby B and Kara K (2011) Spatial prediction of AADT in unmeasured locations by universal kriging, transportation research board 90th annual meeting
19. Wu M and Tae JKK (2022) A hybrid geostatistical method for estimating citywide traffic volumes: a case study of Edmonton, Canada. *J Geogr Res* 5(02)
20. Aljamal MA, Abdelghaffar HM, Rakha HA (2020) Real-time estimation of vehicle counts on signalized intersection approaches using probe vehicle data. *IEEE Trans Intell Transp Syst* 22(5):2719–2729
21. Dorigo M, Birattari M, Stutzle T (2006) Ant colony optimization. *IEEE Comput Intell Mag* 1(4):28–39
22. Li Y, Soleimani H, Zohal M (2019) An improved ant colony optimization algorithm for the multi-depot green vehicle routing problem with multiple objectives. *J Clean Prod* 227:1161–1172
23. Zhou R, Nee A, Lee H (2009) Performance of an ant colony optimization algorithm in dynamic job shop scheduling problems. *Int J Prod Res* 47(11):2903–2920
24. Gao W, Tang Q, Ye B, Yang Y, Yao J (2020) An enhanced heuristic ant colony optimization for mobile robot path planning. *Soft Comput* 24(8):6139–6150
25. Behrisch M, Bieker L, Erdmann J, and Krajzewicz D (2011) Sumo-simulation of urban mobility: an overview, SIMUL 2011 the third international conference on advances in system simulation
26. Arthur D and Vassilvitskii S (2007) K-Means++: the advantages of careful seeding, proceedings of the 18th annual ACM-SIAM symposium on discrete algorithms, SODA '07: 1027–1035