IEICE TRANS. COMMUN., VOL.E104–B, NO.7 JULY 2021

PAPER *Special Section on Future Directions of Research and Development on Communication Quality*

Efficient Data Diffusion and Elimination Control Method for Spatio-Temporal Data Retention System*

Shumpei YAMASAKI[†a)], ***Student Member***, **Daiki NOBAYASHI**[††b)], **Kazuya TSUKAMOTO**[†††c)],
Takeshi IKENAGA[††d)], ***Members, and*** **Myung J. LEE**[††††e)], ***Nonmember***

SUMMARY With the development and spread of Internet of Things technologies, various types of data for IoT applications can be generated anywhere and at any time. Among such data, there are data that depend heavily on generation time and location. We define these data as spatio-temporal data (STD). In previous studies, we proposed a STD retention system using vehicular networks to achieve the "Local production and consumption of STD" paradigm. The system can quickly provide STD for users within a specific location by retaining the STD within the area. However, this system does not take into account that each type of STD has different requirements for STD retention. In particular, the lifetime of STD and the diffusion time to the entire area directly influence the performance of STD retention. Therefore, we propose an efficient diffusion and elimination control method for retention based on the requirements of STD. The results of simulation evaluation demonstrated that the proposed method can satisfy the requirements of STD, while maintaining a high coverage rate in the area.

key words: *STD, data retention, diffusion and elimination, dynamic transmission interval control*

1. Introduction

The development and progress of machine-to-machine (M2M) communication and the Internet of Things (IoT) are accelerating the emergence of numerous IoT devices and new applications. According to Cisco Annual Internet Report (2018–2023) [1], global M2M connections will grow 2.4 times from 6.1 billion in 2018 to 14.7 billion by 2023. Besides, connected car is expected to be the fastest growing IoT application type at 30% CAGR over the forecast period (2018–2023). Due to this trend, a large amount of data generated by IoT devices (e.g., sensors) are expected to flow through networks, and thus a large number of services using the sensor information will appear. For processing and distributing these data, network equipment with higher performance will be required. In addition, high-performance platforms are needed to analyze and leverage vast amounts of sensor data for more productive applications. Since cloud systems, one of the technologies supporting the current IoT era, are centralized systems, the load on network and computing equipment will increase as data grows. On the other hand, some data generated by IoT devices depend on location and time, such as traffic, weather, disaster, and time-limited store advertisement, and we define such data as spatio-temporal data (STD).

With existing network architectures, STD are collected in remote servers or clouds connected to the Internet, and the servers or clouds process STD as a part of IoT application data so that they can be used by users. Along this line, although users need to search and request appropriate STD when necessary, it is not practical especially for the STD with strict requirements. For example, we consider a case where car accident (i.e., one of STD) happens at a certain intersection. In this case, the accident information should be delivered to the drivers of the cars behind the road as soon as possible. However, in the cloud system, the information firstly should be collected by the cloud servers and then the information and/or related information such as traffic jam somehow should be delivered to the eligible drivers as soon as possible. That is, the method of STD collection and delivery with the minimum latency is essential. In this regard, since the devices such as nearby traffic lights and vehicles exist in the vicinity of where the STD is generated. STD can be directly collected by these devices and then delivered to the users from these devices. Therefore, in this paper, we focus on the location-aware system by utilizing neighboring devices for "Local production and consumption of STD."

M. Beuchert et al. [2] proposed a platform that makes it easy to manage and analyze STD for location-based applications, but it does not consider delivering STD in real-time. K. Tsukamoto et al. [3] proposed a Geolocation-Centric Information Platform (GCIP) that consist of hybrid between edge network and ad-hoc network. The GCIP delivers STD that is generated in a location and timely-aware manner. In this paper, the authors also considered the case where the edge network such as base station and Wi-Fi AP is broken

---

Manuscript received September 1, 2020.
Manuscript revised November 27, 2020.
Manuscript publicized January 8, 2021.

[†] The author is with the Kyushu Institute of Technology (KIT), Iizuka-shi, 820-8502 Japan.

[††] The authors are with the Department of Electrical and Electronics Engineering, Faculty of Engineering, Kyushu Institute of Technology, Kitakyushu-shi, 804-8550 Japan.

[†††] The author is with the Department of Computer Science and Networks, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, Iizuka-shi, 820-8502 Japan.

[††††] The author is with the Department of Electrical and Computer Engineering, City University of New York, USA.

[*] An earlier version of the present paper was presented at the 11th International Workshop on Information Network Design (WIND-2019).

a) E-mail: p232087s@mail.kyutech.jp
b) E-mail: nova@ecs.kyutech.ac.jp
c) E-mail: tsukamoto@cse.kyutech.ac.jp
d) E-mail: ike@ecs.kyutech.ac.jp
e) E-mail: mlee@ccny.cuny.edu

DOI: 10.1587/transcom.2020CQP0010

due to some reason. In this case, both the collection and delivery of STD are maintained by the ad-hoc network (not edge network) in the simplicity manner.

Therefore, we need to propose a system that delivers STD at specific time and location without infrastructure-based edge network. The system has a potential for suppressing the load on core network as much as possible. Furthermore, resilient STD collection and delivery functions can be maintained only by the ad-hoc networks. That is, the paradigm of “Local production and consumption of STD” has a potential to be achieved without the STD search method in the ad-hoc manner.

H. Teshiba et al. [4] proposed the spatio-temporal data retention system (STD-RS) using pure vehicular ad-hoc networks (VANET) to achieve the paradigm of “Local Production and Consumption of STD” without the support of existing network infrastructure. This system focuses on vehicle’s mobility and the possibility that vehicles are commonly equipped with a high-performance CPU and radio communication I/F. This system treats a vehicle as an information hub (InfoHub) that relays information and is responsible for diffusing and maintaining STD in a specific area. In this system, InfoHubs adaptively “retain” (i.e., periodically broadcast) STD within the area to achieve the floating state of STD. This brings users in the specific area to passively receive STD by InfoHubs in a real time manner. In summary, the system provides the following advantages:

- STD retrieval in a passive way (no need to search/request STD)
- Real-time STD collection and distribution
- Providing fault tolerance
- Providing data offloading from the infrastructure network

In addition, previous studies [4], [5] proposed a control method of data transmission probability based on vehicle density in order to avoid frame collisions with neighboring vehicles. However, although each type of STD has different requirements in terms of its lifetime and its diffusion completion time, previous methods did not consider these requirements. In this paper, we consider to diffuse and retain the wireless resource information such as available wireless bandwidths/channels. Since the information deeply depends on a specific time and area (i.e., spatio-temporal characteristics), wireless resource information can be treated as a one of STD. Since this STD consisting of the wireless resources tend to change in the real time manner, prompt diffusion is essential for efficient utilization of the STD. Therefore, the STD retention performance are suffering from the following two issues.

1. Some STD may not fully be spread over an entire specific area by the time users need it. This results in the degradation of STD retention performance.
2. STD may be unnecessarily remained even beyond the validity period determined by the STD origin. This results in the redundant resource consumptions.

As a result, STD retention could be ineffective in our previous studies.

In the present paper, we propose an efficient data diffusion and elimination control method for the STD-RS. First, we define the target time of data diffusion completion. Our goal is that STD be reliably diffused in an entire specific area by the target time of diffusion completion, while suppressing the number of data transmissions and packet losses. To maintain the consistent control throughout from data diffusion to data retention, we herein focus on the data transmission interval, which directly affects the data diffusion speed, and thus propose a new method in which the data transmission interval is dynamically changed for quick STD diffusion. Furthermore, to avoid the redundant resource consumptions, we also propose a STD elimination method in which vehicles immediately eliminate the STD when their lifetime has expired. The effectiveness of the proposed method is evaluated through simulation experiments.

The remainder of the present paper is organized as follows. Related research is reviewed in Sect. 2. In Sect. 3, we describe the STD-RS proposed in previous research, and the details of the two proposed methods (STD diffusion control and the STD elimination control) are explained in Sect. 4. Section 5 presents and discusses the performance evaluation. Finally, concluding remarks and future research are discussed in Sect. 6.

2. Related Work

F. Li et al. [6] discussed the problem of VANET, such as data diffusion and data sharing, and proposed a protocol that uses Geocast Routing based on location within a specific area and sends information from a source to all nodes (Vehicles).

C. Maihofer et al. [7] proposed a method by which data are delivered to vehicles within the retention area and are stored for a certain period. Furthermore, they proposed three methods for keeping/delivering a geocast message depending on the location of the retention area: (1) the server approaches, (2) the election approaches, and (3) the neighbor approaches.

First, in the server approaches, a fixed server performs data retention and transmits data to the retention area based on the Geocast Routing protocol. The server needs to exchange the location information of each node, which may increase the network load. Moreover, A. Maio et al. [8] proposed a method for setting an optimum retention area range using a software-defined network (SDN). The server plays the role of the SDN controller, decides the optimum retention area radius from the movement information of the neighboring nodes, and the data transmissions are performed for data retention. Second, in the election approaches, the selected node stores the data and sends the data to the retention area. These two systems impose a heavy load on a specific server/node and may not be able to maintain data distribution due to the failure of a specific server/node.

Finally, the neighbor approaches use a system consisting of only mobile nodes. G. Rizzo et al. [9] proposed a practical application for exchanging and sharing disaster information among mobile nodes without using network infrastructure. In addition, Refs. [5] and [10], Floating Content [11], Locus [12], etc., have been proposed as a system consisting of only mobile nodes. In Ref. [10], a node heading to a retention area is specified by exchanging the navigation information of each node, and data are effectively transmitted. In the Floating Content and Locus systems, each node has a list of maintaining data, and exchanges the data list with neighboring nodes. When a node does not have the data, the node sends a transmission request to the neighboring node to obtain the data. Since whether a node performs transmission is determined by the transmission probability according to the distance from the center, data acquisition decreases when the node is far from the center, whereas data collision may occur frequently when the node is biased to near the center.

In order to solve the problems of the related researches, we have proposed the STD-RS in which the STD is periodically broadcast by appropriate vehicles within the retention area. As a result, all users within the area can passively receive the STD in the STD's transmission interval. In the next section, the technique devised in the construction of the STD-RS is explained.

3. Spatio-Temporal Data Retention System (STD-RS)

In this section, we first describe the objective and target application of STD-RS in Sect. 3.1. In Sects. 3.2 and 3.3, we indicate the assumptions and requirements of the system proposed in our previous study. Then, we explain the control method of data transmission probability proposed in that study [5] in Sect. 3.4.

3.1 Objective and Target Application of STD Retention

The objective of the STD-RS is to spread and then retain STD over the entire retention area. As a result, all users located within a certain area will be able to receive the STD from at least one node of surrounding vehicles in a passive way (without contents request). Furthermore, the system can provide the distributed management of STD without relying on the existing network infrastructure, thereby not only reducing server load effectively but also improving the system's fault tolerance.

Next, we show our target application of the STD-RS. We assume that wireless resource information such as available wireless bandwidths/channels are treated as STD in specific time and area. Users within the specified time and area can passively collect STD from neighboring nodes and then can reliably use an available bandwidth/channel satisfying their communication requirements, without relying on existing network infrastructure.

Since the STD including the wireless resource information has stronger spatio-temporal characteristics (changes tend to occur frequently), the STD retention system should be consisted from the following three steps.

Phase 1: Quick STD diffusion in a whole specific area by using vehicles (STD diffusion).

Phase 2: Periodic STD retention at the specific area by using vehicles (STD retention).

Phase 3: Quick STD elimination after the expiration of STD lifetime (STD elimination).

Our previous work [4], [5] proposed the methods for STD retention (Phase 2). We describe the concrete assumptions, requirements, and the method for STD retention in the next section.

3.2 Assumptions

Originally, the STD-RS is supposed to retain STD in any place and any shape of area. However, in this paper, the STD-RS assumes that the STD are retained continuously in the *retention area* without routing, where the *retention area* denotes a circle of radius R [m] centered on a fixed information source. Figure 1 shows the configuration of the system. We define the radio range as r, and assume that the r of every vehicle is the same. Hereinafter, "vehicle" is referred as to "node." We assume that each node is aware of its position information by using GPS. Each node broadcasts a beacon containing its ID at regular intervals b [s]. By broadcasting beacons, each node can inform other neighboring nodes of its presence, while each node can obtain the number of surrounding nodes. Note that we do not include GPS information of each node in the beacon because the increase of the beacon size reduces the efficiency of channel utilization. In addition, each node independently broadcasts STD to retain them in the retention area. We assume that the STD include the information of the retention area (location of the information source and radius of retention area R) and the data transmission interval d [s], which d is a fixed value determined by the information source in order to retain STD. After receiving STD, each node calculates the distance from the location of the information source included in the received STD using its position information, and each node judges whether it is in the retention area. If each node is not in the retention area, the nodes remove the STD quickly. In contrast, the nodes make preparations for STD transmission when they are in the area.

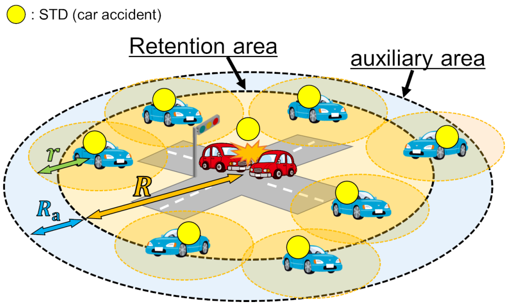


Fig. 1 Configuration of the STD-RS.

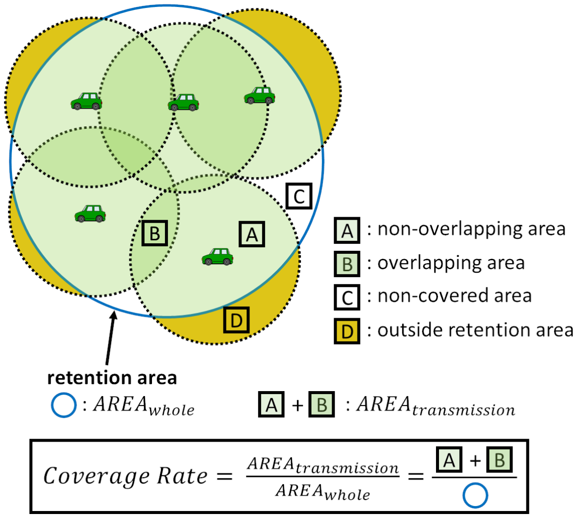


Fig. 2 How to calculate *Coverage Rate*.

3.3 System Requirements

Here, we define a requirement of the system as *Coverage Rate*. The coverage rate indicates the probability that system users can automatically receive STD when entering the retention area:

$$\mathit{Coverage\ Rate} = \frac{\mathit{AREA}_{transmission}}{\mathit{AREA}_{whole}}, \tag{1}$$

where $AREA_{transmission}$ is the total area where users located in the retention area can receive STD within the STD transmission interval d, and $AREA_{whole}$ is the entire retention area. Figure 2 shows how to calculate *Coverage Rate*. The blue circle represents the retention area whose size is defined as $AREA_{whole}$. On the other hand, the green color spaces represent the area where users can receive STD in the retention area. That is, the sum of the area A and B is defined as $AREA_{transmission}$. However, since the users located in the area C and D cannot receive the STD, the space of the area C and D are not included in the $AREA_{transmission}$. Since STD is retained within the retention area for a certain period continuously, the *Coverage Rate* requires to make itself high rate every data transmission interval d.

The previous study proposed a transmission probability control method based on node density in order to suppress useless data transmission. Section 3.4 shows how to control transmission probability.

3.4 Data Transmission Probability Control Based on Node Density

In this system, we define the auxiliary area, as shown in Fig. 1. The width of the auxiliary area is defined as R_a. Then, each node decides whether to transmit the STD based on the distance between their location and the location of the information source (defined as *Distance*). The transmission target area is defined as follows:

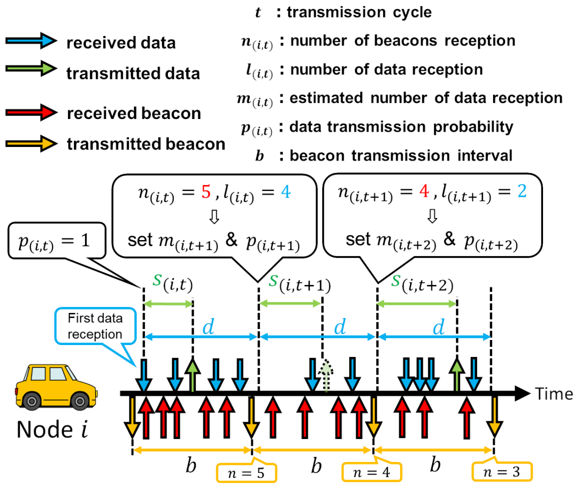


Fig. 3 Setting data transmission probability.

$$\begin{cases} 0 < \mathit{Distance} \le R + R_a. & : \text{transmission} \\ \mathit{otherwise}. & : \text{non-transmission} \end{cases} \tag{2}$$

All nodes within the retention area can directly contribute to STD retention. However, since the number of nodes near the border of the retention area is inherently decreased, the coverage rate may be degraded especially in case of low node density environment. Therefore, we ask the nodes within the area between R and $R + R_a$ (i.e., outer edge of the retention area) to contribute the STD retention. Note that we already confirmed that the size of R_a should be set to at least 75% of the radio coverage r.

Next, we describe the STD transmission timing. Figure 3 shows the operation of STD and beacons transmission by a node i, where i is a unique ID given to the node. First, node i counts the number of received beacons during each of the beacon transmission intervals (b seconds) to estimate the number of neighboring nodes n. From this, each node can detect the node density within its own radio range.

Second, we describe the transmission operation of STD. We assume that the periodical transmission cycle (data transmission interval d) is determined by the information source and included in the STD. Therefore, the node i can know the predetermined STD transmission cycle whenever receiving the STD. However, if multiple nodes receive the same STD simultaneously, the transmission cycles overlap and data collisions may frequently occur. To solve these issues, every node sets the STD transmission timing for every transmission cycles as at random time $s_{(i,t)}$ $(0 < s \le d)$ in order to avoid data collisions, where t is a sequence number of the transmission cycle.

Lastly, we introduced stochastic STD transmission control. If all nodes in the retention area broadcast the STD in every d, data collisions prevent effective STD retention. Furthermore, even if we want to achieve a high coverage rate, it is redundant for multiple nodes to send STD to the same part of the retention area over and over again. Therefore, in this system, the transmission probability $p_{(i,t)}$ is controlled based on the neighboring node density and the number of STD transmission of the surrounding nodes as following steps:

Step 1: When the node i initially receives a STD from other nodes, the transmission probability $p_{(i,1)}$ is set to 1 because the STD may not be diffused in the whole retention area. If the node already has the STD, move to **Step 2**.

Step 2: Each node decides the transmission probability $p_{(i,t)}$ based on the number of neighboring nodes $n_{(i,t)}$. Here, when the number of nodes is more than four, the node's transmission range has the potential to be covered by that of the neighboring nodes. For example, when the neighboring four nodes are located to its north, south, west, and south (ideal arrangement), the node's transmission cover area is completely enclosed by that of other nodes. Therefore, when the number of neighboring nodes is three or less, a node sets $p_{(i,t)}$ to 1 because the node needs to cover the transmission range itself. On the other hand, when the number of neighboring nodes is four or more, $p_{(i,t)}$ is determined according to the number of STD transmissions by neighboring nodes in order to prevent excessive STD transmissions and STD collisions. We proceed the **Step 3** to estimate the number of STD transmissions by neighboring nodes.

Step 3: Each node calculates the estimated number of STD transmissions by neighboring nodes $m_{(i,t)}$ as following (3).

$$m_{(i,t)} = \alpha \times l_{(i,t-1)} + (1 - \alpha) \times m_{(i,t-1)}, \tag{3}$$

where $l_{(i,t-1)}$ is the number of received STD in the previous cycle, and α is the moving average coefficient. We continue to the **Step 4** to calculate the transmission probability.

Step 4: Each node calculates the transmission probability based on the $m_{(i,t)}$. Here, we define the target value of the number of STD transmission by neighboring nodes in the next cycle as β. In order to cover the node's transmission area while decreasing the number of STD transmissions in the next cycle, each node decides the transmission probability so that the number of STD transmissions approaches β. The transmission probability is calculated as following (4):

$$p_{(i,t)} = \begin{cases} p_{(i,t-1)} + \frac{\beta - l_{(i,t-1)}}{n_{(i,t-1)}+1} & (0 < m_{(i,t)} < \beta) \\ p_{(i,t-1)} & (m_{(i,t)} = \beta) \\ p_{(i,t-1)} - \frac{l_{(i,t-1)} - \beta}{n_{(i,t-1)}+1} & (\beta < m_{(i,t)}) \end{cases} . \tag{4}$$

The STD transmission probability control is continuously performed while a STD is retained.

In this paper, we use this probability control method of STD transmission based on node density in order to maintain the retention after the STD are fully spread for the entire retention area.

3.5 Problems of the Previous Study

The previous study set the data transmission interval d so that nodes deliver STD to users. In other words, users can certainly receive the STD if they wait as long as up to d seconds for the STD in the retention area. However, the previous study did not assume a concrete constraint on the data transmission interval d and did not mention how to set the data transmission interval.

Some STD with strict constraints, such as a short lifetime and diffusion completion time, will be retained, as well as other STD with loose constraints (e.g., weather or advertising information). For example, these STD include radio resource information, such as available frequency bands and channel interference information. If the transmission interval of STD is longer than its lifetime, i.e., $d >$ [lifetime of STD], then the STD cannot be completely diffused within the area until the expiration of the lifetime (i.e., ineffective retention), thereby some users cannot receive the STD within its lifetime. Therefore, these STD need the transmission interval d to be set as $0 < d <$ [lifetime of STD]. However, even if d is set so as to satisfy $0 < d <$ [lifetime of STD], the users may not be able to fully utilize STD. For instance, if STD diffusion ends just before the STD lifetime expires, the users who received the STD just before the expiration of the lifetime cannot use the STD sufficiently. Furthermore, if STD are retained after the STD lifetime expires, the subsequent STD transmissions are treated as the unnecessary overhead.

Thus, when we handle STD retention with a severe time constraint, we need to consider STD transmission and elimination based on the required time for STD diffusion completion and the lifetime of the STD. Therefore, we herein propose a method by which to control the data transmission interval and eliminate the STD in order to achieve efficient STD retention for the STD.

4. Proposed Methods

In this section, we describe our proposed method for quickly diffusing and eliminating STD. Note that this paper focuses on how to (1) diffuse the STD before the STD retention and to (2) eliminate the STD after the retention period, while our previous work achieved the STD retention by the STD transmission control method we explained in Sect. 3.4.

4.1 Objectives and Policy of the Proposed Method

The objective of our proposed method is that users fully use STD in their valid period. In order to achieve this, we must not only diffuse STD in the entire retention area by target time but also delete STD after the lifetime of the STD.

Although our previous study focused solely on STD retention, we herein introduce STD diffusion and elimination. Therefore, the STD retention is considered to consist of three phases:

1. Diffusion phase
2. Retention phase
3. Elimination phase

Figure 4 shows the transmission phase transition.

Step 1: First, the information source transmits STD which include the parameters for STD diffusion and retention, such

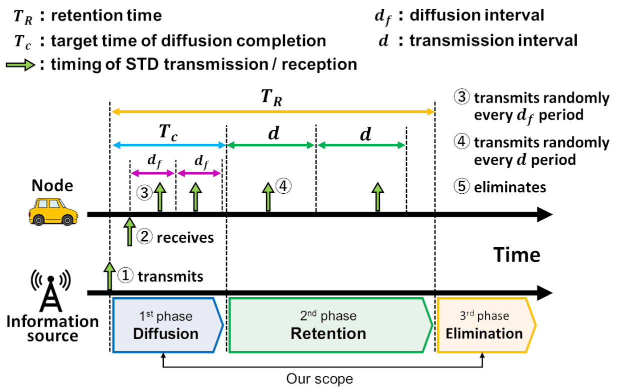


Fig. 4 Data transmission interval during retention time.

as the radius of retention area R, the data transmission interval d, and so on.

Step 2: The node receives the STD, and checks the parameters.

Step 3: Each node diffuses STD in the entire retention area by target time (**Diffusion phase**).

Step 4: Next, they retain STD by the lifetime of the STD after the target time (**Retention phase**).

Step 5: Lastly, immediately after the lifetime of the STD expires, the nodes eliminate the STD in order to avoid redundant STD retention (**Elimination phase**).

In this paper, we consider the first (diffusion) and third (elimination) phases.

First, we define the lifetime of STD, which is the valid period of the STD, as the retention time T_R. We also define the target time of data diffusion completion as T_c. Since the requirements for T_c differ depending on the type of STD, the possible range of T_c could be $0 < T_c \leq T_R$. For instance, the wireless frequency resource can be treated as a STD. Since the availability of frequency resources (bands) is varied in the spatio-temporal manner, the STD retention time T_R tends to become short, and thus T_c also needs to be set to shorter value.

Next, we need to consider how to achieve STD diffusion completion by T_c. Although multi-hop communication is considered as one of the methods for STD diffusion, the STD-RS does not need any kinds of routing protocols, that is, none of addition overhead is introduced. In addition, the system aims to diffuse STD quickly, while maintaining the data transmission probability control in the consistent manner. Note that we aim to realize STD diffusion just by changing parameters of data transmission probability control.

Therefore, we focus on that the time of STD diffusion completion varies by adjusting the value of STD transmission interval d. In the diffusion phase, we modify the d for STD diffusion, which is determined based on the number of neighboring nodes and the size of the retention area. We define the modified data transmission interval for STD diffusion before the STD retention as d_f ($< d$). In diffusion phase, each node transmits STD every d_f in order to quickly diffuse STD within the whole retention area as shown in Fig. 4. After T_c seconds pass (diffusion phase is over), each node sets the transmission interval to d longer than d_f in order to limit the unnecessary data transmissions as much as possible during the retention phase.

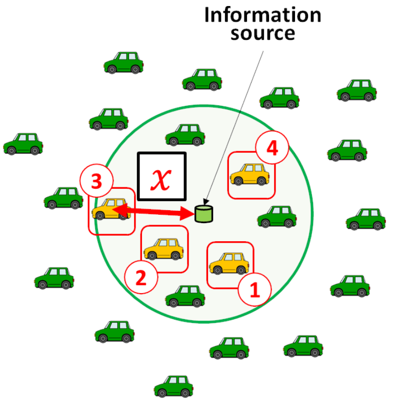


Fig. 5 Example in which the nodes sequentially perform STD transmission ($\gamma = 4$).

In this paper, we propose a method in which the information source dynamically decides d_f based on the number of neighboring nodes n and the target time of diffusion completion T_c, as shown in Fig. 4. The procedures of STD diffusion are as follows. First, an information source detecting and having some STD decides the retention area R, the retention time T_R, and the target time of diffusion completion T_c. Second, to diffuse the STD to the whole retention area R, the information source determines the data transmission interval for STD diffusion d_f just before transmitting the STD, based on the number of neighboring nodes n, which is estimated by the number of beacons received in the last b seconds before the calculation of d_f, where b is the beacon transmission interval. Intermediate nodes receiving the STD randomly determine the transmission timing within the period of d_f.

4.2 Transmission Interval Control for Data Diffusion d_f

First, we define the minimum number of nodes to diffuse STD in all directions as γ. Figure 5 shows an example in which the nodes sequentially perform STD transmissions after an information source transmits STD in case of $\gamma = 4$. The information source is located at the center of the retention area, and the green circle indicates the radio range in which the information source can diffuse STD. When the information source transmits a STD, the seven nodes in the green circle can receive the STD. Each node first checks the transmission interval d_f, which is predetermined by the information source and included in the received STD, and decides the transmission timing s_j ($0 < s_j \leq d_f$) randomly, where j denotes an order in which STD transmission of neighboring node ($\{j \mid 1 \leq j \leq n,\ j \in \mathbb{N}\}$). The s_j indicates that the estimated time until a j-th node out of the n nodes completes its transmission. Since each node determine s_j by random number following the uniform distribution, the cumulative distribution function value of the j-th transmission timing becomes $\frac{j}{n}$. As a result, the estimated time s_j can be expressed as $\frac{j}{n} \times d_f$. Next, we assume that the STD are diffused in all directions when the STD transmission is performed γ times even if all seven nodes do not perform STD transmission. At this time, we define the expected time in which the γ-th STD transmission is completed as t_γ. The t_γ can be expressed by the following (5):

$$t_\gamma = \frac{\gamma}{n} \times d_f. \tag{5}$$

Here, since the time at which one hop transmission ends should be t_γ, the number of transfer hops n_h required before T_c can be calculated by (6):

$$n_h = \frac{T_c}{t_\gamma} = \frac{T_c}{\frac{\gamma}{n} \times d_f} = \frac{nT_c}{\gamma d_f}. \tag{6}$$

Furthermore, we define the distance from the information source to the farthest among the γ nodes as x, as shown in Fig. 5 (in this case, $\gamma = 4$). As a further requirement, we need to complete STD diffusion in the entire retention area by the n_h hops transmission. In order to achieve the requirement, $n_h \times x$ must exceed the circle of radius R. Therefore, d_f is derived as follows:

$$\begin{aligned} x \times n_h &= R \\ x \times \frac{nT_c}{\gamma d_f} &= R \\ d_f &= \frac{n}{\gamma} T_c \times \frac{x}{R}. \end{aligned} \tag{7}$$

Since each node randomly determines the STD transmission timing s within the d_f, the diffusion distance x is changed according to the location and density of neighboring nodes. That is, x cannot be known by the information source in advance. Hence, we consider the following three cases for estimating the distance of x.

1. O_case: Optimistic case
2. E_case: Expected value case
3. P_case: Pessimistic case

(1) O_case is a state in which nodes exist on the edge of the radio coverage r of some node. In the case of this node arrangement, since the STD are always spread with the radius of r by one STD transmission. As a result, the required number of hops that can cover the retention area becomes the minimum value.

(2) E_case is a case where the diffusion distance x is determined by the expected value. In this case, we assume that all nodes are deployed uniformly in the radio coverage r. First, we divide the circle of radius r into an area of x $(< r)$ radius circle (*area A*) and an outside area (*area B*). Being that x is the expected value is the same thing as that the number of nodes of *area A* and *B* are equal. As a result, the expected value of x is the radius of *area A*. That is, x becomes $\frac{1}{\sqrt{2}}r$.

(3) P_case is defined as the distance $\frac{r}{2}$, which is smaller than the expected value. In this case, the required number of hops becomes double that of the optimistic case.

The values of d_f in the O_case, the E_case, and the P_case are shown in d_{f_o}, d_{f_e} and d_{f_p} of following (8), respectively.

$$\begin{cases} d_{f\_o} = \frac{n}{\gamma} T_c \times \frac{r}{R}. & (x = r) \\ d_{f\_e} = \frac{1}{\sqrt{2}} \frac{n}{\gamma} T_c \times \frac{r}{R}. & (x = \frac{1}{\sqrt{2}} r) \\ d_{f\_p} = \frac{1}{2} \frac{n}{\gamma} T_c \times \frac{r}{R}. & (x = \frac{1}{2} r) \end{cases} \tag{8}$$

The relative relation of the distance x is O_case > E_case > P_case. That is, the initial transmission intervals d_f has a relationship of d_{f_o} (O_case) > d_{f_e} (E_case) > d_{f_p} (P_case). Therefore, it is assumed that O_case and P_case can diffuse STD within the retention area up to T_c in case of the node densities are high and low, respectively. On the other hand, it is assumed that E_case can satisfy the requirement of T_c irrespective of node density.

4.3 Elimination Procedure

If the STD retention is still maintained after the expiration of its retention time T_R, then the STD retention waste not only the wireless but also storage resources. In order to solve this problem, in the proposed method, each node checks the start of STD transmission (T_s) and the lifetime (T_R) included in the received STD. Then, if $T_n - T_s > T_R$, where T_n denotes the present time, then the STD are surely discarded, thereby preventing unnecessary STD distribution. The phase is called elimination phase as shown in Fig. 4.

5. Performance Evaluation

In this section, we evaluate the effectiveness of the proposed dynamic transmission interval control method through network simulation. We first describe the simulation model in Sect. 5.1. In Sect. 5.2, we explain the evaluation indexes and the comparison methods, and then present the simulation results and discussions in Sect. 5.3.

5.1 Simulation Model

We evaluate the proposed method using the network simulator *OMNeT++* [13], the traffic simulator *SUMO* [14], and *Veins* [15], which is the simulation platform implementing IEEE 802.11p.

We assume a grid-shaped road network, as shown in Fig. 6. In addition, we focus on the environment where the location of the information source is fixed at the center of the retention area whose shape is circle. The lane interval is 200 m. The speed of the nodes is set to 40 km/h. Moreover, we prepare two models of vehicular mobility: the *simple model* (Fig. 6) and the *random model*, which was not used in our previous paper [16]. We create the *simple model* and *random model* using *SUMO* simulator [14]. We use the *simple model* to demonstrate whether the proposed method can satisfy the requirements of STD: its diffusion completion time T_c and its retention time (lifetime) T_R. In the *simple model*, nodes move from right to left or

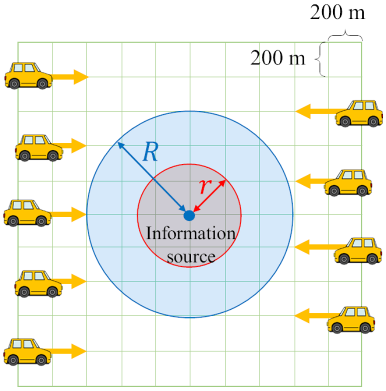


Fig. 6 Simple mobility model.

from left to right (without turns) while keeping the distance between each vehicle constant, which results in the uniformity of node density regardless of location. We use four node density environments of the *simple models* in which the average number of neighboring nodes is 2.5, 3, 5, 7, 9, and 14. On the other hand, the *random model* is used to show that the proposed method can operate in a real environment. In the *random model*, nodes decide start point and end point, and they move the shortest distance between the two points. Although movement pattern is multiple, one route is selected randomly. We vary the number of average neighboring nodes as 4, 6, 8, 10, 12, 14, 16, 18, and 20. In addition, we use five random mobility patterns for each average number of neighboring nodes. Furthermore, we set the radio communication range r of nodes to 300 m. The minimum number of nodes to diffuse in all direction γ is set to 4 based on our previous study. In this paper, we assume that STD-RS diffuses and retains the STD that frequently changes in a short period, such as wireless resource information. Therefore, we set short value for the target time of diffusion completion T_c and retention time T_R. In addition, in terms of the requirements of STD, we set the retention radius R, the retention time T_R, and the target time of diffusion completion T_c to 600 m, 4 s, and 1 s, respectively. The T_R is set so short in order to evaluate STD elimination method this time. Each node transmits beacons every 1 s, and the packet size of the STD is fixed to 300 bytes. In present paper, we do not discuss the effect of the packet size. An information source transmits STD ten times for each mobility pattern.

5.2 Evaluation Indexes and Comparison Methods

First, we here introduce three performance indexes employed in this paper. First one is **Coverage Rate at the time of** T_c. Since the objective of the proposed method is to diffuse the STD over the whole retention area up to the diffusion completion time T_c, we employ the coverage rate at the time of T_c as the performance index. Second index is the total **number of STD transmissions up to the** T_c. The proposed method tries to maximize the Coverage Rate at the time of T_c, while reducing the number of STD transmissions up to the time of T_c as much as possible. Third index is the **number of STD loss up to the** T_c. The definition of "STD loss" is the cumulative number of STDs that vehicles cannot receive due to radio interference. Hence, the increase in the number of STD transmissions inevitably causes the increase in the STD loss, thereby resulting in insufficient STD distribution and/or retention.

Next, we explain the two comparative methods to indicate the effectiveness of our proposed method. As mentioned above, the objective of this paper is to diffuse STD in the whole retention area until T_c. Therefore, we first adopt the flooding method, which is a common method for diffusing messages in VANET and MANET, as a comparative method. Since each node re-broadcasts data whenever receiving messages in the flooding method, the flooding method can maximize the coverage rate at T_c. However, the number of data transmissions drastically increases with the increase in the number of nodes in the retention area, which results in the increase in the STD losses. Hence, it is important whether our proposed method can suppress the number of STD transmissions and their losses, while achieving the coverage rate at T_c equivalent to that of the flooding method. The second comparative method is our previous method proposed by reference [4]. Since the main purpose of the previous method is retention of STD, the method randomly determined the transmission timing within the transmission interval d (not using the diffusion interval d_f). That is, the previous method did not focus on the STD diffusion, and thus the diffusion cannot be reliably completed up to the T_c. On the other hand, in this paper, since we treat the STD having the strong spatio-temporal characteristics (i.e., the T_R becomes short), the role of STD diffusion we focus on in this paper becomes crucial. Note that parameters employed for the previous method such as d and T_R are set based on the reference [4] (d of 5 seconds and T_R of 10 seconds, which is twice of d). Other parameters of the previous method are set to the same as those of our proposed method and flooding methods.

5.3 Simulation Results and Discussions

In Sect. 5.3.1, we demonstrate whether the proposed method can operate properly while satisfying the requirements of STD in the *simple model*. Then, in Sect. 5.3.2, we show whether the proposed method can achieve the requirements of STD in the model of random vehicular mobility, which simulates a real environment.

5.3.1 Evaluation in Model of Simple Vehicular Mobility

Figure 7 shows the average coverage rates and the maximum and minimum values with the change in the average number of neighboring nodes. From these results, every method is difficult to diffuse STD to the whole retention area up to T_c, when the average number of neighboring nodes is 2.5 and 3. Note that the limit of STD retention method with the low-density environment in the "retention" phase is proposed and evaluated in [17]. Next, we focus on the

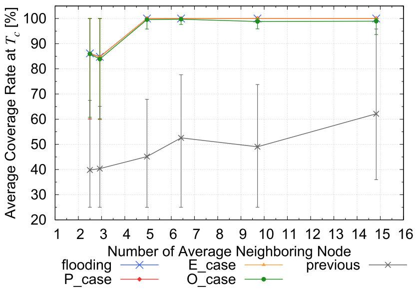


Fig. 7 Average coverage rate at T_c.

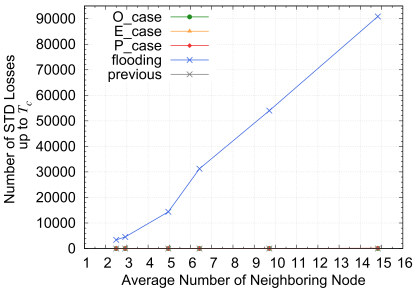


Fig. 9 Number of STD losses up to T_c.

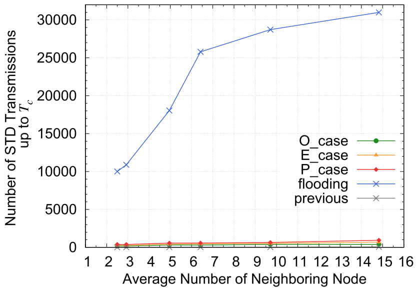


Fig. 8 Number of STD transmissions up to T_c.

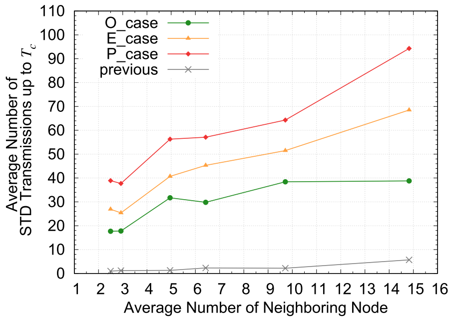


Fig. 10 Average number of STD transmissions up to T_c (proposed method).

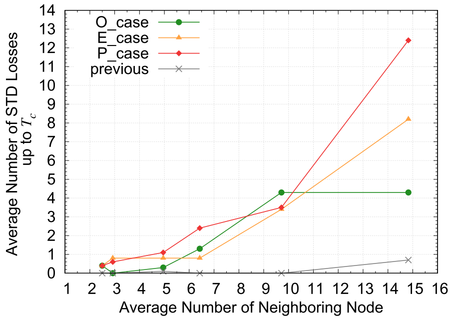


Fig. 11 Average number of STD losses up to T_c (proposed method).

case where the number of neighboring nodes is more than 5. The previous method cannot provide approximately 100% of the average coverage rate at T_c, regardless of the number of neighboring nodes. Since the previous method did not concern about the STD diffusion and thus each node relatively slowly broadcasts the STD based on the d, the STD cannot be diffused to the whole retention area at the T_c (= 1 second). If we consider the STD with relatively short validity time, such as the available wireless resource information, the delay of completion of STD diffusion may not utilize the STD efficiently due to the prompt change in their availability. On the other hand, both the flooding method and our proposed method achieves an average coverage rate of 95% or more. However, the difference between the maximum and minimum values in O_case is the largest compared to other methods. Since O_case assumes that nodes exist right on the radio coverage, the result indicates that O_case cannot flexibly adapt to the location of nodes.

Figures 8 and 9 show the number of STD transmissions up to T_c and the number of STD losses, respectively. As shown in the figures, both values of flooding method are quite large. It is self-evident that the values for flooding method increase steadily as the number of neighboring nodes increases. Although the flooding method may be the quickest method to diffuse STD in the entire retention area, the high number of STD transmissions does not indicate that the method can diffuse STD efficiently. In contrast to flooding method, the proposed method can avoid STD losses by reducing the number of STD transmissions while maintaining a high average coverage rate at T_c.

Figures 10 and 11 show the average number of STD transmissions up to T_c and the average number of STD losses up to T_c obtained by previous method and each of the proposed methods, respectively. First, in the previous method, both the average number of STD transmissions and their losses up to T_c become the lowest value compared with those of other methods. However, the achievable coverage rate at T_c becomes a significantly lower value due to the fewer number of STD transmissions. Next, we focus on our proposed method. O_case always maintains the smallest number of STD transmissions, whereas P_case always keeps the largest values of the three methods, irrespective of the average number of neighboring nodes. Since O_case assumes that the neighboring nodes are located on the border line of transmission coverage only and calculates the d_f based on the assumption, the method does not flexibly adapt to the changes in the number of neighboring nodes, thereby resulting in the fluctuation of the coverage rate at T_c. On the other hand, since the average number of STD transmissions of P_case becomes the largest, the average number of STD losses tends to be the highest with some variation. Since the proposed method aims to achieve not only reduction of the number of STD transmissions but also maintenance of a high coverage rate, we can see that the E_case is an appropriate method.

5.3.2 Evaluation in Model of Random Vehicular Mobility

Figures 12, 13, and 14 show the median, the maximum, and minimum values of the average coverage rate at T_c for each average number of neighboring nodes in the random model.

From these results, when the average number of neighboring nodes is smaller than seven, the coverage rates of almost every method do not reach 90% or more. Moreover, the coverage rates are dynamically fluctuated, because the number of nodes diffusing STD is not sufficient, or nodes temporally do not exist uniformly even if a sufficient number of nodes exist in the retention area.

Focusing on the part where the average number of neighboring nodes is seven or more, P_case (Fig. 14) always achieves a coverage rate of 90% or more. E_case (Fig. 13) fails to reach 90% coverage rate when the number of neighboring nodes is 8.3 and 18.5, otherwise reaching 90% or more.

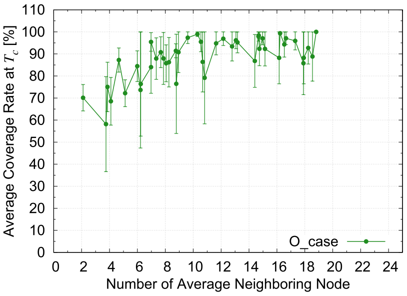


Fig. 12 Median coverage rate at T_c (O_case).

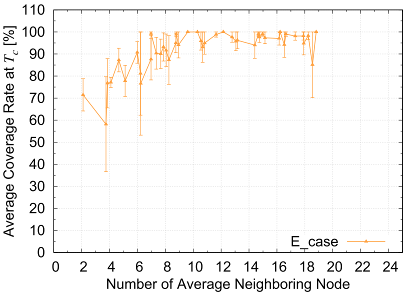


Fig. 13 Median coverage rate at T_c (E_case).

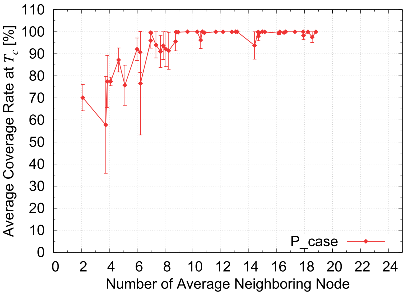


Fig. 14 Median coverage rate at T_c (P_case).

However, O_case sometimes suffers the decrease in the coverage rate (less than 90%), thereby providing unstable performance. Furthermore, O_case assumes that γ nodes exist on the radio coverage of an information source and calculates d_f. As a result, if the γ nodes exist at a shorter distance than the radio coverage of the InfoHub vehicles, then d_f is set to be longer than the transmission interval that achieves a high coverage rate up to T_c.

Figures 15 and 16 show the number of STD transmissions up to T_c and STD losses of the proposed method with the change in the average number of neighboring nodes. Regardless of the average number of neighboring nodes, the number of STD transmissions and the number of STD losses tend to be the same relationship of O_case < E_case < P_case because of the predetermined transmission interval of each method. These results indicate that O_case is the most efficient method in terms of the STD transmissions in the diffusion phase. However, O_case provides unstable coverage rate and suffers from the coverage of less than 90%, as we mentioned before. As a result, E_case can effectively reduce the number of both STD transmissions and losses while maintaining a high coverage rate.

5.3.3 Evaluation of STD Elimination Mechanism

So far, we evaluated the effectiveness of our proposed method for STD diffusion. In this section, we evaluate the effectiveness of the STD elimination method. Figure 17 shows the distance from the information source at which the

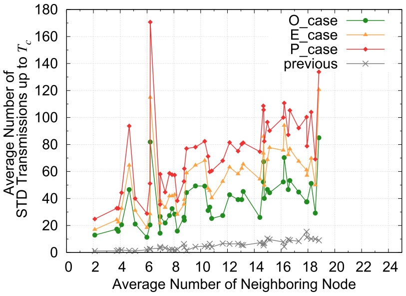


Fig. 15 Number of STD transmissions up to T_c.

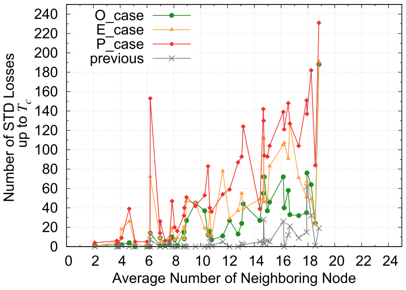


Fig. 16 Number of STD Losses up to T_c.

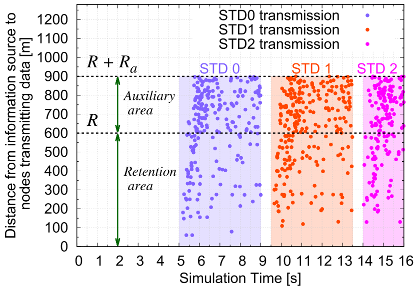


Fig. 17 Time and location of STD transmissions.

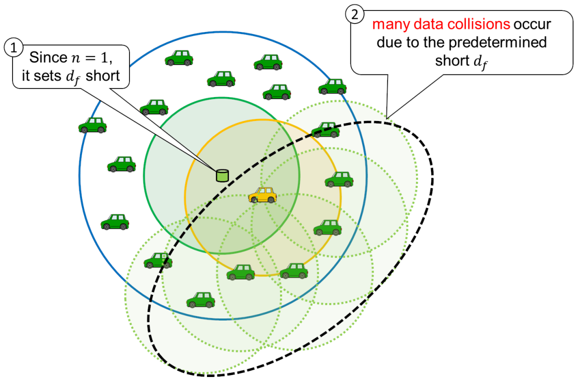


Fig. 18 Dynamic adjustment of d_f.

STD transmissions took place with the time series. We here employ the random mobility model for vehicle movement and conduct simulations in the severe environment where the average number of neighboring nodes is set to 20 (high density) and the P_case method with the shortest diffusion interval is employed. Ten different STD (STD0, STD1, ..., STD9) are retained during one simulation trial. However, we present the records of only first three STD retained from simulation start up to 16 s. Note that in this simulation evaluation, we set the retention time T_R to 4 s. As shown in Fig. 17, we can see that each STD is surely retained during its retention time T_R. In other words, we were certainly able to eliminate these STD once their retention time T_R completes and avoid permanently retaining STD.

5.3.4 Discussion of the Future Work

In Sects. 5.3.1, 5.3.2, and 5.3.3, we demonstrated the effectiveness of the proposed method. However, in the proposed method, the STD transmission interval in the diffusion phase, d_f, is determined based on the neighboring nodes of the information source only. That is, the InfoHub vehicles cannot adjust the d_f according to the change in the surrounding environment. This may result in an inefficient STD distribution depending on the surrounding environment. For instance, in a case where the number of neighboring nodes of an information source is small, but there are many nodes within the retention area, the predetermined d_f could not be appropriate for efficient STD retention. As illustrated in Fig. 18, if the number of neighboring nodes n is one, the information source sets d_f to be a short value by following the proposed scheme. As a result, the large number of nodes receiving the STD are enforced to transmit STD every predetermined short d_f, thereby experiencing the large number of STD transmissions and losses. To solve this problem, each intermediate node needs to autonomously update the d_f according to the change in its surrounding node density. Therefore, we will tackle a problem how to dynamically update the d_f by the intermediate nodes for providing efficient STD retention, in the near future. Moreover, we need to verify the proposed method using an actual mobility model that more closely approximates the real environment.

6. Conclusion

In the present paper, we focused on the STD retention system as a way of achieving the paradigm of "Local production and consumption of STD." Although the previous method considered the STD transmission probability, a method by which to diffuse the STD in the initial phase and to eliminate STD that expire in the third phase has not been considered. Hence, in the present paper, we proposed a dynamic transmission interval control method to complete STD diffusion by the target time, and a STD elimination method to eliminate STD exceeding the STD's lifetime. Through simulations, we showed that the proposed scheme, especially E_case, can reliably achieve a high coverage rate, while significantly limiting the increase in the number of STD transmissions. Moreover, we can demonstrate that the STD retention system can reliably eliminate STD in response to the STD lifetime.

In the future, we will consider how each intermediate InfoHub node dynamically updates the STD transmission interval according to the change in their surrounding environment. In addition, we intend to verify the proposed method using an actual mobility model that more closely approximates the real environment so that the proposed method can work in any vehicle environment.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 18H03234, NICT Grant Number 19304, and USA Grant number NSF 17-586.

References

[1] Cisco, "Cisco Annual Internet Report (2018–2023) White Paper," https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf, 2020.
[2] M. Beuchert, S.H. Jensen, O.A. Sheikh-Omar, M.B. Svendsen, and B. Yang, "aSTEP: Aau's spatio-temporal data analytics platform," 19th IEEE International Conference on Mobile Data Management (MDM), pp.278–279, Aalborg, Denmark, 2018. DOI: 10.1109/MDM.2018.00049
[3] K. Tsukamoto, H. Tamura, Y. Taenaka, H. Yamamoto, T. Ikenaga, and M. Lee, "Geolocation-centric information platform for resilient spatio-temporal content management," IEICE Trans. Com-

mun., vol.E104-B, no.3, pp.199–209, March 2021.

[4] H. Teshiba, D. Nobayashi, K. Tsukamoto, and T. Ikenaga, "Adaptive data transmission control for reliable and efficient spatio-temporal data retention by vehicles," 16th International Conference on Networks, pp.46–52, Venice, Italy, 2017. DOI:10.1109/CloudNet.2018.8549521

[5] T. Higuchi, R. Onishi, O. Altintas, D. Nobayashi, T. Ikenaga, and K. Tsukamoto, "Regional InfoHubs by vehicles: Balancing spatio-temporal coverage and network load," Proc. IoV-VoI'16, pp.25–30, 2016. DOI:10.1145/2938681.2938688

[6] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," IEEE Veh. Technol. Mag., vol.2, no.2, pp.12–22, 2007. DOI:10.1109/MVT.2007.912927

[7] C. Maihofer, T. Leinumuller, and E. Schoch, "Abiding geocast: Time-stable geocast for ad hoc networks," Proc. 2nd ACM VANET'05, pp.20–29, 2005. DOI:10.1145/1080754.1080758

[8] A. Maio, R. Soua, M. Palattella, T. Engel, and G. Rizzo, "A centralized approach for setting floating content parameters in VANETs," 14th IEEE Annual CCNC, pp.712–715, Las Vegas, USA, Jan. 2017. DOI:10.1109/CCNC.2017.7983220

[9] G. Rizzo and H. Neukirchen, "Geo-based content sharing for disaster relief applications," International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, Advance in Intelligent System and Computing, vol.612, pp.894–903, Torino, Italy, July, 2017. DOI:10.1007/978-3-319-61542-4_91

[10] I. Leontiadis, P. Costa, and C. Mascolo, "Persistent content based information dissemination in hybrid vehicular networks," Proc. IEEE PerCom, pp.1–10, Texas, USA, 2009. DOI:10.1109/PERCOM.2009.4912758

[11] J. Ott, E. Hyyti, P. Lassila, T. Vaegs, and J. Kangasharju, "Floating content: Information sharing in urban areas," Proc. IEEE PerCom, pp.136–146, Seattle, USA, 2011. DOI:10.1109/PERCOM.2011.5767578

[12] N. Thompson, R. Crepaldi, and R. Kravets, "Locus: A location-based data overlay for disruption-tolerant networks," Proc. ACM CHANTS, pp.47–54, New York, USA, 2010. DOI:10.1145/1859934.1859945

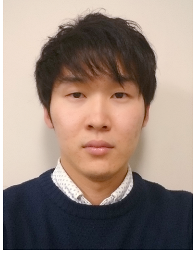
[13] OMNeT++, "OMNeT++ discrete event simulator," OMNeT++, https://omnetpp.org, accessed Dec. 29. 2019.

[14] SUMO, "SUMO - Documentation," Institute of Transportation Systems, https://sumo.dlr.de/docs/SUMO_User_Documentation.html, accessed Dec. 29. 2019.

[15] C. Sommer, "Veins - The open source vehicular network simulation framework," Veins, http://veins.car2x.org, accessed Dec. 29. 2019.

[16] S. Yamasaki, D. Nobayashi, K. Tsukamoto, T. Ikenaga, and M. J. Lee, "On-demand transmission interval control method for Spatio-Temporal data retention," The 11th International Workshop on Information Network Design (WIND-2019), vol.1035, pp.319–330, Oita, Japan, 2019. DOI:10.1007/978-3-030-29035-1_31

[17] I. Goto, D. Nobayashi, K. Tsukamoto, T. Ikenaga, and M.J. Lee, "Transmission control method to realize efficient data retention in low vehicle density environments," INCoS 2019, AISC 1035, pp.390–401, Sept. 2019.



Shumpei Yamasaki received the B.E. degree from Kyushu Institute of Technology, Iizuka, Japan in 2019. Since April 2019, he has been a Master cource student at Kyushu Institute of Technology. His recent research interests are an information distribution using VANET. He is a student member of IEICE.

Daiki Nobayashi received the B.E., M.E. and D.E. degrees from Kyushu Institute of Technology, Tobata, Japan in 2006, 2008, and 2011, respectively. From April 2010 to March 2012, he was a Japan Society for the Promotion of Science (JSPS) Research Fellow. From 2012 to 2021, he was an assistant professor at the Department of Electrical and Electronics Engineering, Faculty of Engineering, Kyushu Institute of Technology. Since April 2021, he has been an associate professor at the Department of Electrical and Electronics Engineering, Kyushu Institute of Technology. His research interests include network architectures and protocols, wireless LANs, multi-hop wireless networks, wireless sensor and actuator networks. He is a member of IEEE and IEICE.

Kazuya Tsukamoto received the DE degree in computer science from the Kyushu Institute of Technology (KIT), Japan, in 2006. From April 2006 to March 2007, he was a Japan Society for the Promotion of Science (JSPS) research fellow at KIT. In 2007, he was an assistant professor in the Department of Computer Science and Electronics, KIT, and then has been an associate professor in the same department since April 2013. His research interests include performance evaluation of computer networks and wireless networks. He is a member of the ACM, the IPSJ, the IEICE, and the IEEE.

Takeshi Ikenaga received B.E., M.E. and D.E. degrees in computer science from Kyushu Institute of Technology, Iizuka, Japan in 1992, 1994 and 2003, respectively. From 1994 to 1996, he worked at NEC Corporation. From 1996 to 1999, he was an Assistant Professor in the Information Science Center, Nagasaki University. From 1999 to 2004, he was an Assistant Professor in the Department of Computer Science and Electronics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology. Since September 2011, he has been a Professor in the Department of Electrical, Electronic Engineering, Faculty of Engineering, Kyushu Institute of Technology. His research interests include performance evaluation of computer networks and QoS routing. He is a member of IEEE and IEICE.

Myung J. Lee received BS, MS from Seoul National University and Ph.D from Columbia University respectively. He is currently a professor with the Dept. of Electrical and Computer Engineering, City University of New York. His current researches include Secure V2X communications, Edge Cloud resource management, IoT, machine learning, stochastic computing application to intrusion detection. He publishes extensively in these areas and hold over 25 U.S. and international patents. His research group developed the first NS-2 simulator for IEEE 802.15.4, a standard NS-2 distribution widely used for wireless sensor network research. Dr. Lee co-received three best paper awards. He actively contributed to international standard organizations IEEE (TG chairs for 802.15.5 WPAN Mesh and 802.15.8 PAC) and ZigBee. He is a past president of the KSEA.