Adaptive Data Transmission Control for Reliable and Efficient Spatio-Temporal Data Retention by Vehicles

Hiroki Teshiba*, Daiki Nobayashi†, Kazuya Tsukamoto† and Takeshi Ikenaga†
* Graduate School of Computer Science and System Engineering, Kyushu Institute of Technology
† Network Design Research Center, Kyushu Institute of Technology
680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan
Email: *teshiba@infonet.cse.kyutech.ac.jp, †{nova, tsukamoto, ike}@ndrc.kyutech.ac.jp

Abstract—**Vehicles, which have penetrated deeply into society and have become essential to our daily lives, have two major characteristics that are comparable with conventional communication devices (such as cellular phones). First, since each modern vehicle is now equipped with large computational power and vast data storage capacity, they can easily collect, process, and individually store vast amounts of data. Second, they have remarkably high mobility, and can thus transport and spread stored data to everywhere very effectively. Therefore, in this study, we focus on vehicular ad-hoc networks (VANETs) constructed solely with vehicles, and without any support from outside infrastructure. On one hand, although users can usually receive various application services through the Internet, some specific services, such as those handling traffic and local weather information, are strongly dependent on geographical location and time (and so this information is referred to as spatio-temporal data in this paper), which is not readily available via the Internet. Therefore, as a means to providing spatio-temporal data reliably and effectively by exploiting VANET, we propose an adaptive transmission control method in which each vehicle controls data transmission probability by considering the data retention density of neighboring vehicles. Through simulations, we found that our proposed method is effective for retaining spatio-temporal data.**

Keywords*–*VANET, Data retention, Adaptive data transmission control

I. INTRODUCTION

With the progress and widespread dissemination of machine-to-machine (M2M) and Internet of Things (IoT) technologies, the number and types of devices equipped with various wireless modules have expanded rapidly. In current Internet paradigms, most data are first gathered into remote servers connected to networks, after which they are provided to applications as required. However, according to an Organisation for Economic Co-operation and Development (OECD) report [1], the number of M2M devices will grow to fifty billion by 2020, and enormous amounts of small data will flow to the Internet. In order to store and process these data effectively, the acquisition of large-capacity storage modules and high-performance central processing units (CPUs) is essential.

From the viewpoint of data contents, some specific applications such as weather and traffic information are strongly dependent on location and time. Therefore, the utilization of data collected from the IoT devices, which are referred to as *spatio-temporal data* in this paper, can be expected to improve the quality and accuracy of such information. Since the “locally produced and consumed” paradigm of spatio-temporal data use is effective for location-dependent applications, a novel network architecture that can achieve data retention within a specific area is crucial.

In this paper, we focus on vehicular ad-hoc networks (VANETs) as an important network infrastructure that can achieve the required level of spatio-temporal data retention. Modern vehicles have two remarkable features. First, data can be collected by and analyzed within individual vehicles because they are now equipped with significant amounts of storage space, battery power, and high-level computational resources. Second, since there are enormous numbers of highly mobile vehicles operating all over the world, they can provide a foundation from which data can be collected and/or distributed efficiently.

Furthermore, the potential for spatio-temporal information communication between vehicles in a VANET allows us to advocate a new promising network infrastructure. In our study, we utilize vehicles with spatio-temporal data as *regional information hubs*, or *InfoHubs*, in order to disseminate spatio-temporal data within some pre-defined area. The spatio-temporal data are finally received by users (not vehicles). Spatio-temporal data management by *InfoHub* vehicles brings us the following advantages:

- Users can obtain the spatio-temporal information quickly.
- Thanks to distributed data management, an acceptable level of fault tolerance can be achieved.
- Internet server loads can be reduced.

If spatio-temporal data are managed in a distributed manner, users can obtain the data from neighboring vehicles, thereby achieving real-time data acquisition. Moreover, if data are replicated among multiple vehicles in advance, spatio-temporal data retention can be maintained even when some vehicles break down. Finally, data management by InfoHub vehicles has the potential to decrease power consumption by Internet-based (cloud) servers.

However, since all vehicles in a VANET generally utilize the same communication channel, frame (data) collisions are inevitable and a certain level of interference is inherent. In networks with large numbers of vehicles (dense traffic environments), each vehicle could suffer multiple and frequent frame collisions, leading to a decline in communication quality. On the other hand, in networks with small numbers of vehicles (sparse traffic environments), each vehicle must accelerate data transmission activity due to the lack of data transmission timing. With these points in mind, it is clear that the use of adaptive data transmission control in response to vehicle density could provide an indispensable component for distributed data management by exploiting the capabilities of InfoHub vehicles.

Accordingly, in this paper, we propose an adaptive data transmission control method in which vehicles adaptively change data transmission probabilities in response to the density of neighboring vehicles in order to maintain spatio-temporal data retention within a pre-defined area, and thus allow area users to efficiently obtain local spatio-temporal data. In our proposed method, each vehicle estimates not only the number of neighboring vehicles based on the number of beacons but also the number of received data based on the number of data received thus far. Then, based on the estimated values, each vehicle dynamically changes data transmission probability in a way that facilitates overall spatio-temporal data retention within the specified area. Through simulation-based evaluations, we clarified that our proposed method can always achieve an acceptable level of data retention within the target area, irrespective of vehicle density changes.

The rest of this paper is organized as follows. In Section II, we review related works. In Section III, we describe our spatio-temporal data retention system. Section IV shows the detailed mechanisms of our proposed method. Section V provides the simulation model and simulation results. Finally, Section VI is our conclusion.

II. RELATED WORK

Fan Li et al. discussed various VANET-related problems such as data dissemination and data sharing caused by the high mobility of vehicles [2] and proposed the *Geocast Routing-based protocol*, which is basically a location-based multicast routing, in order to deliver data from a source vehicle to all other vehicles within the target area. Maihofer et al. [3] proposed an *abiding geocast* in which data are delivered to all vehicles within the target area and then maintained within them during the lifetime of the network. They provided three solutions for retaining the geocast data within the target area: (1) *server approach*, (2) *election approach*, and (3) *neighbor approach*. We will provide an overview of these approaches in the following paragraph.

In the server approach, a pre-defined fixed server within the target area is used to store and periodically transmit data to other vehicles within the target area based on a geocast routing protocol. Since the server sends data and exchanges location information among all vehicles within the target area, it is susceptible to overloading. Should that occur, the server would not be able to effectively communicate with vehicles if many failures appear, thereby degrading its dissemination performance. In the election approach, only the elected vehicles maintain the data and periodically send the data to other vehicles within the target area. In both of these two approaches, broadcasting from a restricted number of vehicles can result in spatio-temporal data retention performance degradation.

Finally, the neighbor approach, which consists of only the vehicles without a dedicated server or elected vehicle, has been actively studied recently due to its high feasibility, and a number of systems such as that of [4], Floating Content [5], Locus [6], and our previous work [7], have been proposed. In the method of [4], a vehicle exchanges navigation information with neighboring vehicles, identifies other vehicles that are moving towards the target area, and then delivers the data to them. In the Floating Content and Locus systems, each vehicle has a list of data and exchanges its list with the lists of other vehicles that it encounters. If any vehicle has data that are not stored in a neighboring vehicle, the neighboring vehicle can acquire the data from the vehicle that has the data. In this situation, the vehicle that has the data decides what data to send based on the transmission probability. The transmission probability changes dynamically depending on the distance from where the data were generated. More specifically, the transmission probability decreases as the vehicle moves away from the center of the target area, thereby indicating that some outlying recipients will be unlikely to receive the data. In contrast, if there are numerous vehicles near the center of the target area, data collisions tend to occur frequently in VANETs because each vehicle attempts to send high transmission probability data at the same time.

Meanwhile, unlike Floating Content and Locus, our previous work [7] aims to deliver data to all vehicles within a target area at set pre-determined intervals, employing a geolocation-based broadcasting method. In this method, the transmission probability for periodical data dissemination is determined based on the “location information of all neighboring vehicles”. Thus, this method needs a complicated calculation by vehicles.

In our research, like [7], we focus on a VANET-based system that disseminates and maintains spatio-temporal data within a target area by adaptively controlling data transmission probability in response to the vehicle density, which is estimated from the number of received data transmissions only. In our proposed method, the decision process of transmission probability is simplified because only the message information are employed without the location information of all vehicles like [7]. More specifically, although existing study [7] requires accurate location information of all vehicles in order to calculate the distance between vehicles, it is quite difficult in terms of computational overhead in a practical environment. Therefore, our proposed method only requires the number of broadcast messages to decide the transmission probability. That is, no complex information (e.g., location information) is required. We refer to our VANET-based system as a *spatio-temporal data retention system*.

III. SPATIO-TEMPORAL DATA RETENTION SYSTEM

In this section, we describe the assumptions behind our spatio-temporal retention system (III-A), the objective of our system design (III-B), and its requirements (III-C).

A. Assumptions

Spatio-temporal data are assumed to have originated at a specific location and have a target area within a predetermined radius. Information related to the data’s origination location and target area are hereafter referred to as the “retention requirement” and are included in the spatio-temporal data by the user generating the data.

Since each vehicle can obtain location information by using its GPS receiver and has a unique ID, it can estimate the number of neighboring vehicles based on the received beacon messages broadcast by the InfoHub vehicles. Note that these InfoHub vehicles are equipped with an on-board wireless interface employing IEEE 802.11p specification. Moreover, each vehicle performs an operation that determines whether it is within the target area.

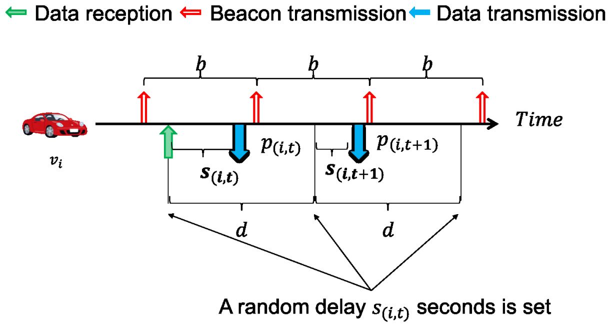


Figure 1. Data transmission procedure.

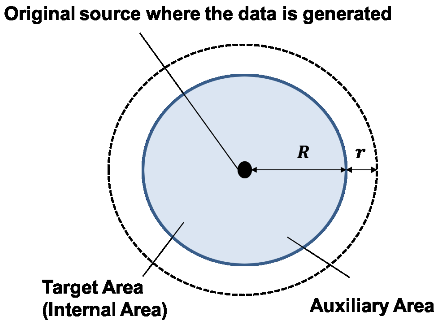


Figure 2. Target area and Auxiliary area.

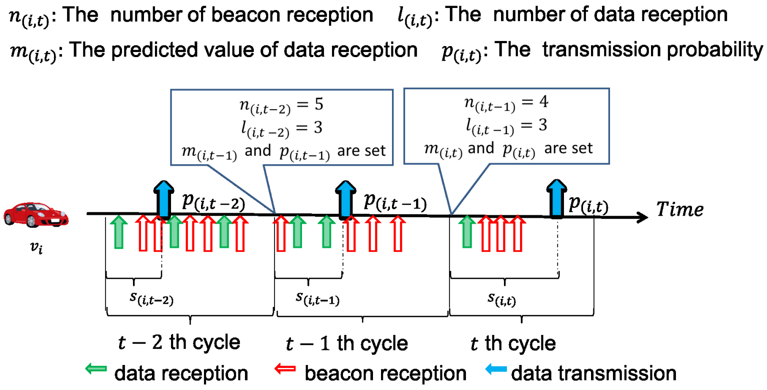


Figure 3. Outline of transmission probability decision.

B. System Objective

The objective of this system is to facilitate data retention, especially for spatio-temporal data such as those on weather and traffic, within a target area. To achieve this, we focus on VANET constructed from vehicles with InfoHub characteristics. By using this system, a user who enters an area can obtain the spatio-temporal data on that area very quickly. Furthermore, since multiple vehicles have the same data, fault tolerance can be achieved. Finally, since the spatio-temporal data are stored only on VANET, there is no burden imposed on Internet (cloud) servers. In the next section, we will discuss the system requirements to achieve our objective.

C. System Requirements

In this paper, we define *coverage rate* as the performance index that indicates how fast users can receive the spatio-temporal data. To facilitate rapid data delivery to users, the entire target area should be covered within the transmission range of InfoHub vehicles. That is, users should be able to obtain the spatio-temporal data from a neighboring vehicle via one-hop broadcast communication. Note that we assume that the transmission range is less than the target area radius, and we calculate the coverage rate at the predetermined interval. The coverage rate formula is shown below:

$$Coverage\ Rate = \frac{S_{DT}}{S_{TA}}$$

where S_{TA} denotes the size of target area, and S_{DT} denotes the size of total area where the user can obtain the data transmitted from either of InfoHub vehicles within the transmission interval. A high coverage rate means that users can automatically receive the spatio-temporal data from anywhere within the target area. Moreover, the slope of the change in the coverage rate indicates the dissemination speed of spatio-temporal data. Therefore, the coverage rate can reveal the responsiveness of the proposed system. Since the proposed system requires rapid acquisition of spatio-temporal data from anywhere within the target area, each vehicle within the area needs to transmit the data as frequently as possible. However, high vehicle density results in frequent data transmissions, which inevitably cause data collisions that can adversely impact the coverage rate. On the other hand, if the vehicle density within the area is low, all vehicles should transmit data as often as possible in order to boost the coverage rate.

In this way, the appropriate transmission probability will change in response to the density of neighboring vehicles. In Section IV, we describe our proposed adaptive transmission control method for achieving the spatio-temporal data retention using InfoHub vehicles.

IV. NODE DENSITY-AWARE TRANSMISSION CONTROL

In this section, we describe our proposed transmission control method, which is based on the number of neighboring vehicles (neighboring vehicle density), and which aims at effective retention of the spatio-temporal data within a target area. Note that, hereafter, InfoHub vehicles are defined as nodes. This method aims to disseminate spatio-temporal data by utilizing the appropriate number of nodes in the target area. Consequently, our spatio-temporal data retention system can maintain a high coverage rate while reducing the total number of data transmissions to the minimum necessary.

A. Data Transmission Timing

In our method, after a node receives data from another node, it needs to re-transmit the received data, as necessary, to ensure spatio-temporal data retention within the target area. However, in order to minimize transmission collisions, the transmission timing of each node is different. This minimizes radio channel collisions among the nodes.

Figure 1 shows the data transmission procedure. In our proposed system, each node periodically transmits the beacon message, but data are only transmitted when necessary. The beacon broadcast interval is fixed at b seconds. On the other hand, the data are transmitted based on the following procedure. When a node v_i receives data from another node, it first checks the transmission intervals of d seconds included in the data. Then, that node randomly determines the next transmission time $s_{(i,t)}$ seconds. Note that the actual transmission time is determined at the start time of the t-th cycle. Here, a cycle lasts d seconds. The random determination within d seconds allows the node to avoid data transmission collisions. This interval d differs between applications, and we assume that the user originating the spatio-temporal data also decides this interval.

B. Adaptive Transmission Control Method

If all nodes are capable of transmitting data at different timing intervals, data collisions can be completely avoided. However, when the number of neighboring nodes within the transmission coverage area is larger than the number of transmission slots, collisions inevitably occur. Accordingly, we designed a new transmission control method in which the transmission probability is dynamically changed based on the neighboring node density, thereby providing a high coverage rate with the minimum number of data transmissions. In our method, nodes around the target area are classified into three types based on distance from the center of the target area (data origin point), as shown in Figure 2. The specific conditions are described below:

$$\begin{cases} 0 \ \le \ x \ \le \ R: & internal\ area \\ R \ < \ x \ \le R + r: & auxiliary\ area \\ otherwise: & independence \end{cases}$$

where x denotes the distance between the node and the center of target area. This distance is calculated from both GPS information and the data origin point, which is included in the data. R shows the radius of the target area, which is referred to as the *internal area*. r is the range of the *auxiliary area*, which is very close to the internal area. The values of R and r are also contained in the data. In Sections IV-B1 and IV-B2, we show how the transmission probability is determined in each area.

1) Internal Area Nodes: The nodes in the internal area autonomously adjust the transmission probability based on the density of neighboring nodes in order to provide a high coverage rate. Figure 3 shows an outline of the transmission probability decision process. The transmission probability $p_{(i,t)}$, which indicates the transmission probability during the t-th cycle, is always set at the start time of the t-th cycle. Note that i represents a unique node ID and t represents a number of cycle.

In the first step, when a node initially receives the data from other nodes, the transmission probability during the first cycle, i.e., $p_{(i,1)}$, is set to 1. That is, the node makes sure to transmit the data because the other nodes cannot provide the data within the receiving node's transmission coverage. This allows us to improve the coverage rate quickly. In the subsequent cycle (t $\geq$2), $p_{(i,t)}$ is determined based on the number of neighboring nodes $n_{(i,t-1)}$. Here, when the number of neighboring nodes is more than four, the node's own transmission range has the potential to be completely covered by that of all neighboring nodes. For example, when the neighboring four nodes are located to its north, south, west, and south (ideal arrangement), the node's potential transmission cover area is already completely enclosed by that of other nodes. Threfore, the decision method of data transmission probability $p_{(i,t)}$ is classified into the following two cases based on the number of neighboring nodes $n_{(i,t-1)}$.

- **case 1** $n_{(i,t-1)} \le 3$:

$p_{(i,t)}$ is set to 1. Since the node's own transmission coverage cannot be completely covered by that of the neighboring nodes, it has to transmit, i.e., $p_{(i,t)}$ is set to 1.

- **case 2** $n_{(i,t-1)} \ge 4$:

$p_{(i,t)}$ is determined based on the number of neighboring nodes and the number of received data. However, since such high node density inherently poses transmission collision risks, only the minimum number of nodes required to maintain the high coverage rate should transmit the data. Conversely, in situations where the location of neighboring nodes is radically asymmetrical and has the potential to become imbalanced, the transmission coverage may not be complete, even if there are a large number of neighboring nodes. This can prevent a node from being able to cover its own transmission range.

To solve these abovementioned problems, we define $m_{(i,t)}$ as the estimated value of the number of received data during t-th cycle and adjust the transmission probability based on the $m_{(i,t)}$. The predicted value $m_{(i,t)}$ is given as equation (1), where $m_{(i,t-1)}$ is the predicted value of the previous cycle, $l_{(i,t-1)}$ is the number of received data in the previous cycle (actual value), and α is the moving average coefficient.

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

The node adjusts its transmission probability so that the number of data transmissions in the t-th cycle becomes the given target value β. If $m_{(i,t)}$ is less than β, the node can predict that the number of data transmissions is likely to be insufficient to cover the area. Therefore, it must increase its transmission probability. On the other hand, if $m_{(i,t)}$ is more than β, the node needs to decrease its transmission probability because excessive data transmissions will occur in the next cycle. At the start of the t-th cycle, each node estimates $m_{(i,t)}$ and then adjusts its transmission probability. Equation (2) describes how the transmission probability is adjusted.

$$p_{(i,t)} = \begin{cases} p_{(i,t-1)} + \frac{\beta - l_{(i,t-1)}}{n_{(i,t-1)}+1} & (0 < m_{(i,t)} < \beta) \\ p_{(i,t-1)} & (m_{(i,t)} = \beta) \\ p_{(i,t-1)} - \frac{l_{(i,t-1)} - \beta}{n_{(i,t-1)}+1} & (m_{(i,t)} > \beta) \end{cases} \tag{2}$$

In this case, the initial value of transmission probability at the first cycle is set to $\frac{\beta}{n_{(i,t-1)}+1}$. This means the average transmission probability of all nodes (including itself and the number of neighboring nodes $n_{(i,t-1)}$) is set to control the number of data transmissions as β. If $m_{(i,t)}$ is less than β, all $n_{(i,t-1)} + 1$ nodes increase their individual data transmission probabilities by $\frac{\beta - l_{(i,t-1)}}{n_{(i,t-1)}+1}$ because their estimates will show that the number of transmitted data does not reach β. On the other hand, if $m_{(i,t)}$ is more than β, the individual nodes

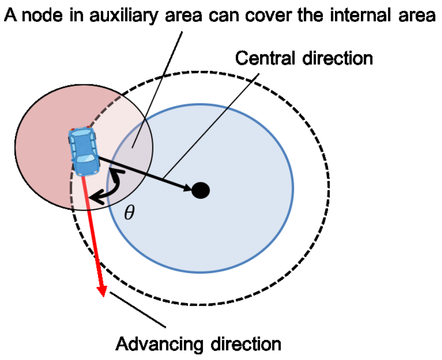


Figure 4. Node behavior in auxiliary area.

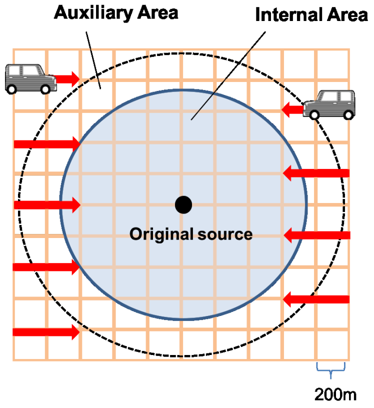


Figure 5. Simulation model.

decrease the transmission probability by $\frac{l_{(i,t-1)}-\beta}{n_{(i,t-1)}+1}$ because they can predict that excessive transmissions will occur. If $m_{(i,t)}$ is equal to β, $p_{(i,t)}$ is set to $p_{(i,t-1)}$ because the current data transmission probability is appropriate. Note that if the value of $\frac{\beta-l_{(i,t-1)}}{n_{(i,t-1)}+1}$ or $\frac{l_{(i,t-1)}-\beta}{n_{(i,t-1)}+1}$ is less than zero, $p_{(i,t)}$ is set to $p_{(i,t-1)}$, and the transmission probability range is varied from $\frac{\beta}{n_{(i,t-1)}+1}$ to 1.

2) Auxiliary Area Nodes: In our proposed method, auxiliary area nodes are also employed to maintain the high coverage rate. To maximize the effect of this extension, just the following two types of nodes are required. The first are nodes that remain in the auxiliary area. The second are nodes approaching the target area (i.e., the angles between the direction of advance and the central direction that are less than θ_{th}, which is the given threshold value, as shown in Figure 4). Since data transmission from these nodes in the auxiliary area can cover the area near the boundary of internal area, these nodes always set $p_{(i,t)}$ to 1. That is, these nodes must transmit the data in order to achieve the high coverage rate.

Finally, the nodes out of the auxiliary area delete the data in order to avoid leaking the spatio-temporal data outside the target area.

V. PERFORMANCE EVALUATION

In this section, we report on a simulation-based performance evaluation of our proposed method. We begin by describing the simulation environment in Subsection V-A. Subsections V-B and V-C present simulation results when the node density and the value of β are changed, respectively. Finally, Subsection V-D shows how the change in node location impacts both the actual number of data transmissions and β. In order to show the effectiveness of our proposed method, we utilized the comparison method called the *naive method*, in which the transmission probability ($p_{(i,t)}$) of all nodes in the simulation area always set to 1.

TABLE I. SIMULATION PARAMETERS.

Internal area	750 m	Auxiliary area	250 m
Transmission range	300 m	α	0.5
Beacon interval	1 s	Transmission interval	5 s

A. Simulation Model

We evaluated our proposed method on the *Veins* [8] simulation platform. The Veins platform implements both the IEEE 802.11p specification for wireless communications and the VANET mobility model, simultaneously. As a result, Veins can combine the network simulator *OMNeT++* [9] and the road traffic simulator *SUMO* [10].

Table I shows our simulation parameters. Here, we assume a grid-shaped road network. The radius of the target area R (distance from the data origination point) is 750 m and the range of the auxiliary area r is set to 250 m, as shown in Figure 5. The velocity of each node on the roads is set to 40 km/h. These nodes run alternately from east to west and from west to east. The communication range of each node is just 300 m. The transmission and the beacon intervals are set to 5 seconds and 1 seconds, respectively.

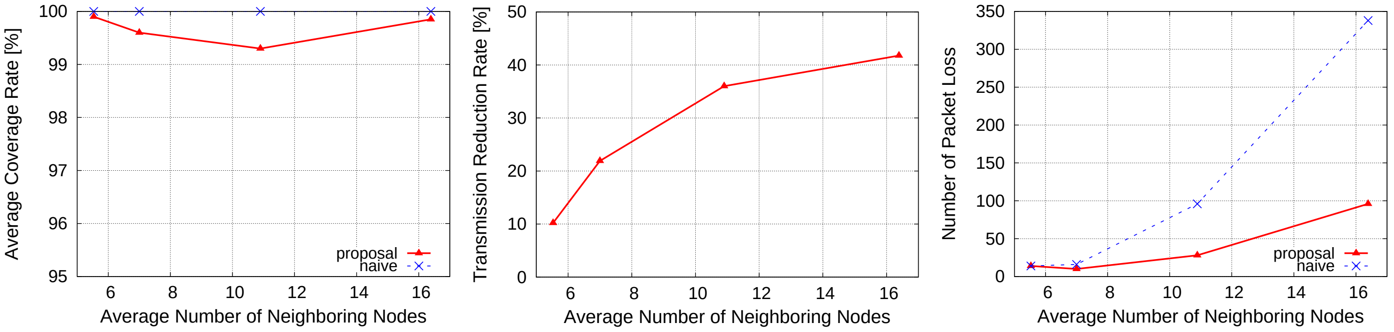
The moving average coefficient α is 0.5. We evaluated our proposed method from the viewpoint of coverage rate and total data transmission reduction over 100 seconds.

B. Node Density Impact

In this subsection, we set $\beta = 4$ because the minimum number of nodes necessary to provide total transmission coverage over the target area is four. In other words, with four nodes available, users can receive data anywhere within the target area. In this environment, we evaluated the performance of our proposed method in cases where the node density changes. The distance between nodes varies from 100 to 300 m. As a result, the average number of nodes within the transmission range of some node (i.e., 300 m) is also varied from approximately 5.5 to 16.4. Therefore, in this subsection, we investigated how node density impacts the coverage rate and the total number of data transmissions.

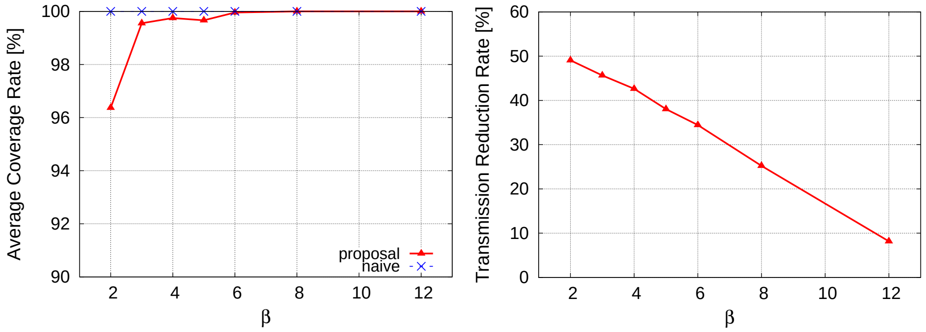
Figure 6 (a) shows the average steady state coverage rate. This steady state denotes a period of 75 to 95 seconds because data retention has already been completed. Since the cycle period is five seconds long, the coverage rate is the average value measured over four cycles (i.e., 20 seconds). From this result, it can be seen that a coverage rate of 99 % or more can be maintained regardless of node density changes.

Figure 6 (b) shows the reduction rate in the total number of data transmissions compared with that of the naive method.



(a) Average coverage rate. (b) Transmission reduction rate. (c) Total number of packets lost.

Figure 6. Performance with varying node density.



(a) Average coverage rate. (b) Transmission reduction rate.

Figure 7. Performance achieved by varying the value of β.

Although the reduction rate is limited to 10 % at the low node density, it can be seen that our proposed method can reduce the total number of data transmissions by 42 % at times of high node density (such as in the case of 16.4 neighboring nodes). This result shows that our proposed method can limit redundant data transmissions effectively as the node density increases. We can also confirm that high transmission probability is set with low node density, whereas low transmission probability is set with high node density.

Furthermore, Figure 6 (c) shows the total number of packets lost when the naive and the proposed method are employed. In particular, it can be seen that, compared with the naive method, our proposal method efficiently reduces the number of total packets lost when the node density is high. From these results, we could confirm that our proposed method can adaptively control the data transmission probability in response to the node density changes while maintaining a coverage rate of approximately 100 %.

C. Impact of the Value of β

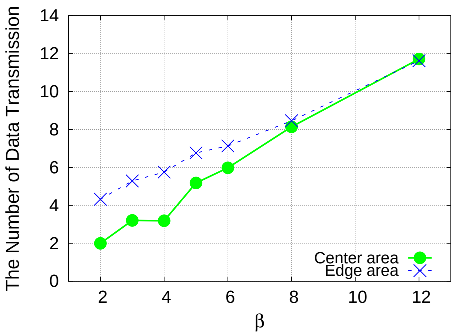
In this subsection, the number of neighboring nodes is fixed at approximate 16.4 and the value of β is varied from 2 to 12. Figure 7 (a) shows the steady state coverage rate with changes in the value of β. This result shows that our proposed method achieves a coverage rate of nearly 100 % except for the case in which β is two. A low β value creates frequent opportunities for data transmission probability decreases, thereby aggressively limiting transmissions. Therefore, the coverage rate cannot reach 100 % if the value of β is low. Figure 7 (b) shows the reduction in the total data transmission rate when the value of β is varied. This result shows that the rate linearly decreases as the value of β increases, and that our proposed method can reduce transmissions by up to 49 % while maintaining a coverage rate of 100 %.

Since the proposed method dynamically changes the data transmission probability in response to location, it is necessary to investigate the change in the transmission probability that occurs with the location consideration discussed in Section V-D.

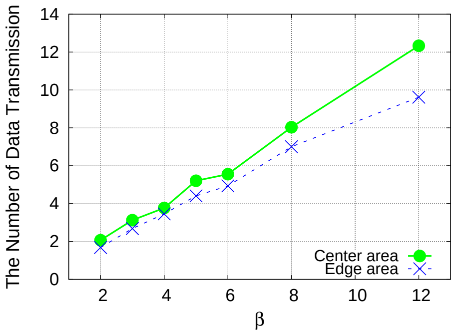
D. Discussion: Location-aware Analysis

In this subsection, we investigate the number of data transmissions that result when the location within the internal area is changed. To achieve this, we separate the internal area into two different sub-areas. (1) *Edge Area*: nodes in this area can receive data transmitted from nodes in the auxiliary area; (2) *Center Area*: nodes in this area do not receive data transmitted from the auxiliary area. The radius of the Center Area is 450 m. The Edge Area is defined as the area within a radius of 750 m but outside the Center Area. We evaluate how the difference of location impacts the number of data transmissions.

Figure 8 (a) shows the average number of data transmissions in Center/Edge Areas. From this result, it can be seen that



(a) Under proposed method Area.



(b) Without data transmissions from auxiliary area.

Figure 8. Average number of data transmissions for nodes in the Center Area and nodes in the Edge Area.

in the Center Area, the number of data transmissions can be controlled to nearly β. However, when the value of β is low, the number of data transmissions in the Edge Area is clearly larger than β. This is because nodes in the Edge Area can receive data from those in the auxiliary area, thereby experiencing many data receptions. Because multiple nodes in the auxiliary area always try to transmit the data, redundant data transmissions occur.

Therefore, to show the contribution of data transmission from nodes in the auxiliary area, we set the probability of those nodes $p_{(i,t)}$ to 0. Figure 8 (b) shows the average number of data transmissions while excluding data transmissions from nodes in the auxiliary area. From this result, we can see that nodes in the Center Area adjust the number of data transmissions to the nearly β. On the other hand, the number of transmitted data from nodes in the Edge Area is insufficient to achieve β, especially in case of high β. This is because the density of nodes in the Edge Area is insufficient and data transmissions from nodes in the auxiliary area are not supported.

From these results, it is clear that the precise control of data transmission from nodes in the auxiliary area is very important for adjusting the number of transmitted data to β. This, in turn, indicates that our proposed method still has an issue that needs to be resolved. Therefore, we will extend our proposed method to permit data transmission probability adjustments for nodes in the auxiliary area, thereby effectively limiting the total number of transmitted data.

VI. CONCLUSION

In this study, our objective was to achieve spatio-temporal data retention within a target area, which would allow users to automatically receive spatio-temporal data from anywhere within the target area. To achieve this, we proposed a new spatio-temporal data retention system that utilizes a VANET constructed from *InfoHub* vehicles. We also proposed an adaptive decision making method for data transmission probability that is based on the density of neighboring vehicles. In our proposed method, each vehicle first estimates the number of neighboring vehicles based on the received beacon messages. Then, the probability is adaptively set with consideration of both the number of neighboring vehicles and the number of transmitted data during the previous time slot. Furthermore, the decision method used differs depending on the vehicle's location (internal or auxiliary area).

Through simulations, we clarified that our proposed method can roughly control data transmissions in response to vehicle density changes. However, we also confirmed that our proposed method still has a problem in which vehicles in the auxiliary area cannot determine an appropriate transmission probability. Thus, in our future work, we will extend the method and then evaluate the improved method under actual traffic environment conditions (such as by using actual traffic data in a real city). Furthermore, although only one type of data is treated in this paper, various types of data would coexist in real environments. Therefore, we will also extend the proposed method in order to treat various types of data simultaneously.

REFERENCES

[1] OECD, “M2M Communications: Connecting Billions of Devices,” *OECD Digital Economy Papers*, No. 192, 2012.

[2] F. Li, and Y. Wang, “Routing in vehicular ad hoc networks: A survey,” *IEEE Vehicular Technology Magazine*, Vol. 2, Issue 2, pp. 12-22, 2007.

[3] C. Maihöfer, T. Leinumüller, and E. Schoch, “Abiding Geocast: Time-stable Geocast for Ad Hoc Networks,” In *Proc. ACM VANET*, pp. 20-29, 2005.

[4] I. Leontiadis, P. Costa, and C. Mascolo, “Persistent content-based information dissemination in hybrid vehicular networks,” In *Proc. IEEE PerCom*, pp. 1-10, 2009.

[5] J. Ott, E. Hyyti, P. Lassila, T. Vaegs, and J. Kangasharju, “Floating Content: Information Sharing in Urban Areas,” In *Proc. IEEE PerCom*, pp. 136-146, 2011.

[6] N. Thompson, R. Crepaldi and R. Kravets, “Locus: A Location-based Data Overlay for Disruption-tolerant Networks,” In *Proc. ACM CHANTS*, pp. 47-54, 2010.

[7] T. Higuchi *et al.*, “Regional InfoHubs by vehicles: balancing spatio-temporal coverage and network load,” In *Proc. IoV-VoI'16*, pp. 25-30, 2016.

[8] “Veins,” [Online]. Available from: http://veins.car2x.org/2017.3.11.

[9] “OMNeT++,” [Online]. Available from: https://omnetpp.org/2017.3.11.

[10] “SUMO,” [Online]. Available from: http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/2017.3.11.