Analysis of Message Delivery Delay in Large-Scale Geographic DTN Routing

Sumika Nishikawa
Department of Informatics
School of Science and Technology
Kwansei Gakuin University
Sanda, Hyogo 669–1337, Japan
E-mail: n-sumika@kwansei.ac.jp

Daiki Matsui, Yasuhiro Yamasaki, and Hiroyuki Ohsaki
Department of Informatics
Graduate School of Science and Technology
Kwansei Gakuin University
Sanda, Hyogo 669–1337, Japan
E-mail: {d-matsui,y-yamasaki,ohsaki}@kwansei.ac.jp

Abstract—**In this paper, we approximately derive the average message delivery delay in geographic DTN routing under random walk mobility on a large-scale network. A geographic DTN routing aims at realization of message delivery among multiple (generally, geographically-dispersed) geographic locations on a field without necessity of specific communication infrastructure by utilizing mobility of mobile agents. In this paper, we address the following research questions. How well or badly does geographic DTN routing perform in a large-scale network (i.e., a network with many geographic locations)? How is the performance of geographic DTN routing affected by the topology of the network (i.e., connections of many geographic locations)? We try to answer the questions using a hybrid modeling of geographic DTN routing with the help of recent advancement in analytical studies of random walks on a graph. We also show that geographic DTN routing is scalable; i.e., its average message delivery delay is approximately proportional to the network size (i.e., geographic locations) unless heavily loaded. We show that the network topology has limited impact on the performance of geographic DTN routing except heavily loaded conditions; the average message delivery delay is mostly determined by the degree of the destination node.**

Keywords*—*DTN (Delay/Disruption-tolerant Networking); Geographic DTN Routing; Multiple Random Walks on Graph; Message Delivery Delay; Large-Scale Network

I. INTRODUCTION

Delay/disruption-tolerant networking (DTN), which allows end-to-end node communication even when communication links between nodes are not functioning normally, has recently been regarded as a promising communication technology for realizing communication infrastructure at the time of disaster and low-cost communication infrastructure [1]. DTN aims to achieve end-to-end data transfers in communication environments where inter-node communication links can be temporarily severed, or the transmission delay between nodes can temporarily increase.

In the literature, extensive researches on DTN routing based on store-carry-and-forward communication with mobile nodes have been actively performed [1]. DTN routing communication between mobile nodes generally utilizes the characteristics of mobile nodes, which autonomously and independently move on the field, and also ad-hoc wireless communication among mobile nodes for realizing end-to-end message delivery.

To the best of our knowledge, most of existing DTN routing algorithms are designed for message delivery between mobile nodes (i.e., message transmission from a mobile node to one or more other mobile nodes) [2-4]. However, in practical applications of DTNs in several fields, endpoints of communication might not be always *mobile* nodes. In other words, endpoints might also be *fixed* nodes, so other types of communications between mobile and fixed nodes and also between fixed nodes should be required. In this paper, a class of DTN routing utilizing mobile nodes for store-carry-and-forward communication among fixed nodes is called *geographic DTN routing* [5, 6].

A geographic DTN routing aims at realization of message delivery among multiple (generally, geographically-dispersed) *geographic locations* on a field without necessity of specific communication infrastructure by utilizing mobility of *mobile agents* [5, 6]. On the field, there exist multiple geographic locations (i.e., fixed nodes) and mobile agents (i.e., mobile nodes), and messages are transferred among geographic locations using store-carry-and-forward operations of mobile agents.

In geographic DTN routing, a message is generated at a geographic location (originating node). Multiple mobile agents independently move on the field. If one of those mobile agents arrives at the originating node, the message is loaded onto the mobile agent. The mobile agent then continues its movement. When the mobile agent arrives at the geographic location to which the message is destined (destination node), the message is unloaded and the destination node receives the message. A single message can be composed of multiple replicas (multiple-copy geographic DTN routing) , rather than a single message (single-copy geographic DTN routing). With multiple-copy geographic DTN routing, the message is considered to be successfully delivered when any of message replicas is delivered to the destination node.

In our previous works, the performance of geographic DTN routing is investigated through simulations [5] and mathematical analysis [6]. Those performance studies show that the geographic DTN routing performs reasonably well. The performance of conventional (i.e., non-geographic) DTN routing is quite limited, in particular, in sparse networks with a low mobile agent density. On the contrary, the performance of geographic DTN routing is favorable since the endpoints of communication are fixed nodes. In other words, the reasonable assumption (i.e., stationary endpoints) significantly boosts the performance of DTN routing.

In this paper, we address the following research questions.

1) *How well or badly does geographic DTN routing perform in a large-scale network (i.e., a network with many geographic locations)?*
   A fundamental question on geographic DTN routing is whether it is scalable in terms of the number of nodes (geographic locations). Note that, in this paper, a *network* means a network of geographic locations, each of which is generally connected with

other geographic locations, rather than a network of mobile agents as in conventional DTN routing. A large-scale network therefore means a network with a large number of geographic locations.

Our previous works [5, 6] investigate the feasibility of geographic DTN routing in rather small-scale networks because of computational complexity of geographic DTN routing simulations and analytical intractability of geographic DTN routing.

Intuitively, the larger the network becomes, the lower the performance of geographic DTN routing becomes since the mobile agent is less likely to arrive at the originating node as well as the destination node. However, it is unclear how quantitatively the performance of geographic DTN routing degrades as the number of geographic locations increases. Understanding the scalability of geographic DTN routing is crucial to design a geographic DTN routing protocol satisfying performance requirements.

2) *How is the performance of geographic DTN routing affected by the topology of the network (i.e., connections of many geographic locations)?*

Another question is on the impact of the network topology on the performance of geographic DTN routing. In this paper, a *network* means a network of geographic locations and the *topology* means the topology of the network composed of geographic locations and connections among them. The performance of geographic DTN routing should be affected by several factors: a geographic DTN routing algorithm, a buffer management mechanism of mobile agents, the capability (e.g., bandwidth and BER (Bit Error Ratio)) of wireless communication among mobile agents and geographic locations, the mobility of mobile agents, and the topology of geographic locations. Among those, the first four factors are *controllable* to some extent. For instance, the capability of wireless communication among mobile agents and geographic locations can be changed by replacing communication protocols and adjusting wireless device parameters. On the other hand, the last two factors are generally *uncontrollable*. It is generally difficult or impossible, for instance, to force mobile agents a specific mobility and/or to change the topology of geographic locations, which usually requires replacement of geographic locations and/or reconstruction of paths among geographic locations. Hence, it is quite important to understand the impact of the network topology on the performance of geographic DTN routing.

In the literature, the impact of the network topology on conventional DTN routing has been investigated [7]. These studies show that the performance of conventional DTN routing is dependent on the underlying network topology. Also, in the field of network science, the relation between the topological structure of a complex graph and its dynamical properties such as the percolation, epidemics, and information dissemination has been extensively studied [8, 9]. These studies show that the network topology considerably affects the dynamical properties such as a probabilistic information dissemination on a complex network. By taking account of these findings, it is natural to assume that the performance of geographic DTN routing should be significantly affected by the network topology since geographic DTN routing has similarity with conventional DTN routing and dynamical processes on a complex network. However, the impact of the network topology on geographic DTN routing has not been well understood.

To the best of our knowledge, one exception is our previous work [6], in which the average and the distribution of message delivery delays are derived. However, as we have explained above, the analytical approach presented in [6] lacks scalability in terms of the network size, which makes it difficult to investigate the impact of the network topology on the performance of geographic DTN routing in medium-scale and large-scale networks.

In this paper, we try to answer the above two questions by extending our previous work — a hybrid modeling of geographic DTN routing [6] — with the help of recent advancement in analytical studies of random walks on a graph. In [6], a hybrid model of geographic DTN routing under random walk mobility on an arbitrary network topology is presented and the major performance metrics of geographic DTN routing such as the average and the distribution of message delivery delays are derived. The key idea in [6] is to combine a continuous-time model (i.e., an $M/M/1$ queueing model) and a discrete-time model (i.e., a discrete multiple random walks on a graph) to model the series of message delivery processes in geographic DTN routing. However, the analytical approach in [6] lacks scalability; the computational complexity to obtain numerical results in [6] grows exponentially as the network size (i.e., the number of geographic locations) increases. So, the analytical approach in [6] is not applicable to answer the above two questions.

The major contributions of this paper are summarized as follows.

- We present a hybrid model of multiple-copy geographic DTN routing, which is composed of a continuous $M/M/1/PS$ queueing model and discrete multiple random walks on a large-scale network
- We approximately derive the average message delivery delay in geographic DTN routing under random walk mobility on a large-scale network
- We analytically obtain the optimal number of message replicas (i.e., the number of copies per a message)
- We show that geographic DTN routing is scalable; i.e., its average message delivery delay is approximately proportional to the network size (i.e., geographic locations) unless heavily loaded
- We show that the network topology has *limited* impact on the performance of geographic DTN routing except heavily loaded conditions; the average message delivery delay is mostly determined by the degree of the destination node

The organization of this paper is as follows. Section II briefly explains the overview and the operation of geographic

DTN routing and a FIFO algorithm. Section III describes our analytic model used throughout this paper. Section IV approximately derives the average message delivery delay using a hybrid model of multiple-copy geographic DTN routing, which is composed of a continuous $M/M/1/PS$ queueing model and discrete multiple random walks on a large-scale network. Section V obtains the optimal number of message replicas that minimizes the average message delivery delay. Section VI presents several numerical examples to address our research questions. Also, the validity of our approximate analysis is investigated by comparing numerical results with simulation ones. Section VII concludes this paper and discusses future works.

II. GEOGRAPHIC DTN ROUTING AND FIFO ALGORITHM

In this section, we briefly explain the overview and the operation of geographic DTN routing. Refer to [5, 6] for details.

A geographic DTN routing aims at realization of message delivery among multiple (generally, geographically-dispersed) *geographic locations* on a field without necessity of specific communication infrastructure by utilizing mobility of *mobile agents*. There exist multiple geographic locations and also multiple mobile agents (i.e., mobile nodes) on the field, and messages are carried by mobile agents for message delivery among geographic locations.

Every geographic location generates messages destined for other geographic locations. We assume that multiple mobile agents autonomously and irregularly visit geographic locations one and another. A mobile agent can *load* a message at its visiting geographic location, carry multiple messages while it moves, and *unload* one or more carrying messages at its visiting geographic location.

There exist a huge number of possible geographic DTN routing algorithms depending on the combination of various factors: message generation patterns and buffer sizes of geographic locations, mobility, mobility controllability, buffer sizes of mobile agents, type and capacity of wireless communications between a geographic location and a mobile agent, and availability of positional information of mobile agents (e.g., GPS (Global Positioning System)).

A typical example of geographic DTN routing is the case where people carrying portable devices such as smartphones autonomously move among geographic locations. Therefore, we assume that mobile agents' mobility are uncontrollable (i.e., a geographic DTN routing algorithm has no control over people's mobility), and the capacity of wireless communication is limited (i.e., the bandwidth for message transfer between a geographic location and a portable device is finite). On the contrary, we assume that the buffer sizes of geographic locations and mobile agents (e.g., portable devices) are sufficiently large.

Geographic DTN routing algorithms can be roughly classified by their message loading mechanism (i.e., how messages are copied/moved from a geographic location) and message unloading mechanism (i.e., how messages are copied/moved from a mobile agent).

In this paper, we focus on one of the simplest algorithms, FIFO (First-In First Out) algorithm, which should be the baseline for other complex geographic DTN routing algorithms.

FIFO algorithm is the minimal and the simplest algorithm, which performs sequential message loading and no message unloading except at the destination node. When a mobile agent visits geographic location v, at most K oldest messages are chosen from the buffer of geographic location v. Those messages are moved to the buffer of the mobile agent. If the mobile agent has one or more messages destined for geographic location v in its buffer, those messages are moved to the buffer of geographic location v.

III. ANALYTIC MODEL

This section presents our analytic model used throughput this paper, which is based on our previous work [6].

We model the field comprising of multiple geographic locations and paths connecting those geographic locations as an undirected graph $G = (V, E)$ where vertices and edges correspond to geographic locations and paths, respectively. Let $d(v)$ be the degree of vertex $v \in V$.

We model the behavior of a mobile agent in geographic DTN routing as a discrete random walk on graph G. At every slot, a mobile agent randomly and synchronously moves one of its neighbor vertices in G. Namely, a mobile agent on vertex v at slot k randomly moves to one of neighbor vertices with probability $1/d(v)$ at slot $k + 1$.

Note that the random walk mobility model is one of the most popular random mobility models in DTN performance studies because of its simplicity and tractability [10]. The random walk mobility model is based on the observation that mobile nodes naturally move around in unpredictable ways [10].

Every mobile agent performs message delivery using FIFO algorithm. When a mobile agent visits a geographic location, it performs the following operations: (1) moves at most K oldest messages from geographic location's buffer to mobile agent's buffer (message loading), and (2) moves, if any, all messages destined for the current geographic location from mobile agent's buffer to geographic location's buffer. The freshness of a message is simply determined by the age of the message (i.e., the time elapsed since the time of its generation). In message loading, if there exist multiple messages of the same freshness, any of those messages are randomly selected. In our analysis, it is assumed that the buffer of geographic locations and mobile agents are sufficiently large.

In this paper, we focus on geographic DTN routing with the number M of mobile agents.

There exist two classes of geographic DTN routing: *single-copy* and *multiple-copy*. In the single-copy case, every message is not duplicated in the network. So, a single-copy geographic DTN routing consumes least network resources. However, message delivery delays in the single-copy case tend to be large, and the message delivery probability from the originating node to the destination node is generally low. For accelerating message delivery and increasing the likelihood of message delivery, a message is duplicated in the multiple-copy case. In this paper, the number of replicas for a message generated at originating geographic location u and destined for geographic location v is denoted by $C_{u,v}$.

We assume that messages are continuously generated at every originating geographic location as a Poisson process. Let $\lambda_{u,v}$ be the message generation rate at originating geographic location u destined for geographic location v. A single message consists of the number $C_{u,v}$ of message replicas. Note that our analytic model is based on discrete random walks on a graph, so our analytic model itself is a discrete model [6].

IV. DERIVATION OF AVERAGE MESSAGE DELIVERY DELAY

In what follows, we derive the average message delivery delay of multiple-copy geographic DTN routing under random walk mobility in a large-scale network.

We focus on message m generated at originating node u and destined for destination node v. For brevity, $C_{u,v}$ is denoted as C; i.e., $C_{u,v}$ and C are used interchangeably. Let $r_1, r_2, \ldots, r_C$ be replicas of message m. In the multiple-copy case, the message delivery delay is defined as the time elapsed since message m is generated at originating node u until one of replicas $r_1, \ldots, r_C$ arrives at destination node v at first.

Let X_i be the random variable representing the delivery delay of replica r_i, which is defined as the duration between the message generation at the originating node and the arrival of replica r_i at the destination node. Also, let X be the random variable representing the message delivery delay. By definition, we have

$$X = \min(X_1, X_2, \ldots, X_C). \tag{1}$$

Hence, the cumulative distribution function of message delivery delay X, $F(X)$, is given by

$$F(x) = 1 - \prod_{i=1}^{C} (1 - F_i(x)), \tag{2}$$

where $F_i(x)$ is the cumulative distribution function of replica delivery delay X_i. The average message delivery delay of message m, $D_{u,v}$, is given by

$$D_{u,v} \equiv E[X] = \int_0^\infty x\, f(x) dx, \tag{3}$$

where $f(x)$ is the probability density function of message delivery delay X (i.e., $f(x) \equiv dF(x)/dx$).

Now we need to derive the cumulative distribution function of replica delivery delay X_i, $F_i(x)$.

The delivery delay of replica r_i is composed of the queueing delay at originating node u and the transfer delay from originating node u to destination node v. More specifically, the delivery delay of replica r_i is the sum of (1) the time elapsed from the generation of message m consisting of all replicas until loading of replica r_i by a mobile agent visiting originating node u (queueing delay), and (2) the time elapsed from the departure of the mobile agent carrying replica r_i until its arrival at destination node v (transfer delay). Random variables representing the queueing delay and the transfer delay of replica r_i are denoted by X_i^Q and X_i^T, respectively. Thus,

$$X_i = X_i^Q + X_i^T. \tag{4}$$

For simplicity, we use the following approximation.

$$X_i \simeq E[X_i^Q] + X_i^T. \tag{5}$$

We then introduce our hybrid model composed of a continuous $M/M/1/PS$ queueing model and discrete multiple random walks on a large-scale network.

First, the departure process from originating node u is modeled as an $M/M/1/PS$ queueing model where messages and loading by a mobile agent in geographic DTN routing correspond to customers and the service in the $M/M/1/PS$ queueing model. Namely, a message, which is composed of the number C of replicas, is regarded as a customer. Once the message (i.e., customer) is generated, it is added to the buffer of originating node u. We assume that when a mobile agent arrives at originating node u, it loads at most the number K of different replicas among all messages queued from originating node's buffer in a round-robin fashion.

Provided that replicas $r_1, r_2, \ldots, r_C$ are processed sequentially and that the departure process from the originating node is stationary, the average queueing delay of replica r_i is approximately given by

$$E[X_i^Q] \simeq \frac{i}{C} E[X_C^Q]. \tag{6}$$

The average queueing delay of the last replica, r_C, is given by the average response time of the $M/M/1/PS$ queueing model:

$$E[X_C^Q] = \frac{1/\mu_u}{1 - \lambda_u/\mu_u}, \tag{7}$$

where λ_u and μ_u are the arrival rate and the service rate of the $M/M/1/PS$ queueing model, respectively. The arrival rate λ_u at originating node u is given by the sum of all message generation rates.

$$\lambda_u = \sum_{v \in V, v \neq u} \lambda_{u,v} \tag{8}$$

Recall that a mobile agent can load at most K replicas at every visit at the originating node. Also recall that every customer (i.e., message) is composed of C replicas. Thus, message m needs C times loading of replicas by mobile agents visiting originating node u. The service rate at originating node u is therefore approximately given by

$$\mu_u \simeq \frac{M\,K}{R_u\,C}, \tag{9}$$

where R_u is the average recurrence time of a mobile agent at originating node u, which is given by the following equation [11].

$$R_u = \frac{2\,|E|}{d(u)} \tag{10}$$

Second, the message transfer with multiple replicas from originating node u to destination node v is modeled as discrete multiple random walks on graph G. All replicas of message m, $r_1, r_2, \ldots, r_C$ are sequentially loaded by different mobile agents visiting at originating node u.

It is shown in [12] that the average hitting time $H_{u,v}$ of a random walk on sufficiently large and reasonably connected graph G starting from vertex u and ending at vertex v is approximately given by

$$H_{u,v} \simeq \frac{2\,|E|}{d(u)}. \tag{11}$$

Assuming that hitting times of a random walk on graph G is exponentially distributed, the probability density function of replica delivery delay X_i is given by the $E[X_i^Q]$-*shifted* exponential distribution with the mean of $H_{u,v}$. For brevity, let Δ be $E[X_i^Q]$.

$$f_i(x) = \begin{cases} \frac{1}{\Delta+H_{u,v}}\, e^{-\frac{x-\Delta}{\Delta+H_{u,v}}} & x \geq \Delta \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Finally, integration of the probability density function $f_i(x)$ yields

$$F_i(x) = 1 - e^{-\frac{x-\Delta}{\Delta+H_{u,v}}}. \tag{13}$$

V. NOTES ON OPTIMAL NUMBER OF MESSAGE REPLICAS

In this section, we analytically obtain the optimal number of message replicas that minimizes the average message delivery delay.

As we will show in Section VI, the optimal number of message replicas in terms of the average message delivery delay is dependent on several system parameters. Among several system parameters, the number of message replicas is easier to change than others such as the number of geographic locations, the topology of the network, and the number of mobile agents. In other words, geographic DTN routing algorithms can dynamically change the number of message replicas to optimize its performance.

One possible application of our approximate analysis is to optimize the number of message replicas at every originating node based on its observations. The optimal number $C^*_{u,v}$ of message replicas for messages originated at geographic location u and destined for geographic location v can be easily obtained by numerically solving the following equation.

$$C^*_{u,v} = \operatorname*{argmin}_{1 \leq C \leq M} D_{u,v} \tag{14}$$

Implementation of such an optimization mechanism in a geographic DTN routing algorithm is not difficult; necessary information for originating node u to obtain the optimal number C^* of message replicas are message generation rate $\lambda_{u,v}$, the maximum number K of message loadings, the number M of mobile agents, the number $|E|$ of edges (i.e., paths) in the network, and the degree $d(v)$ of the destination node. Every geographic location knows the first two, and others are not difficult to obtain or estimate because those are generally not dynamically varying.

VI. NUMERICAL EXAMPLES AND DISCUSSION

In this section, we present several numerical examples to investigate the scalability of geographic DTN routing as well as the impact of the network topology on its performance. Simulation results are also provided to validate our approximate analysis.

For a given network size $N\,(= |V|)$ (i.e., the number of geographic locations), a network topology is synthetically generated using ER (Erdös-Rényi) model [13]. The average degree (i.e., the average number of paths connected to a geographic location) $\overline{k}$ is fixed at $\overline{k} = 6$. Hence, the total number of edges among geographic locations are $|E| = \overline{k}\,N/2 = 3\,N$.

With network topology G of size N, the number of possible originating and destination (i.e., source and sink) node pairs is too many (i.e., N^2) to examine every pair. So, in our experiments, we choose both originating and destination nodes from graph G according to the following rules.

1) The originating node u is set to the node with the largest degree in graph G, which corresponds to the *hub* geographic location in the network.
2) The destination node v is randomly chosen from nodes whose degree is exactly k, which corresponds to a *typical* geographic location in the network.

Note that in our experiments, only a single originating-and-destination node pair is examined.

The message generation rate $\lambda_{u,v}$ has different meanings under different conditions. For instance, if the number M of mobile agents is small and/or the network size N is large, mobile agents are not likely to visit an originating node, resulting in high offered load. For enabling comparison of numerical results under different system parameters, the message generation rate $\lambda_{u,v}$ is normalized using a load factor $\alpha\,(0 \leq \alpha \leq 1)$ as $\lambda_{u,v} = \alpha\,\mu_u$ (see Eq. 9).

Unless stated otherwise, the following system parameters are used: network size (the number of geographic locations) $N = 1,000$, the number of mobile agents $M = 100$, the number of message replicas $C_{u,v} = C = 1$ for all originating and destination node pairs, load factor $\alpha = 0.8$ (i.e., modestly-loaded condition), and the maximum number of message loadings $K = 1$.

Figure 1 shows the average message delivery delay $D_{u,v}$ for different network sizes N. In this figure, the number C of message replicas is changed to 1, 50, or 100. Note that results with $C = 50$ and $C = 100$ (blue and green line) are almost indistinguishable.

Comparison of numerical examples and simulation results is plotted in Fig. 2. Due to computational burden of computer simulations, the maximum network size is limited to $N = 1,000$. This figure shows good agreement between analysis and simulation, in particular, when the number C of message replicas is small. However, we observe some deviation when the number C of message replicas is large.

Figures 1 and 2 seem to indicate that the large number C of message replicas is always desirable regardless of the network size N. However, this is untrue.

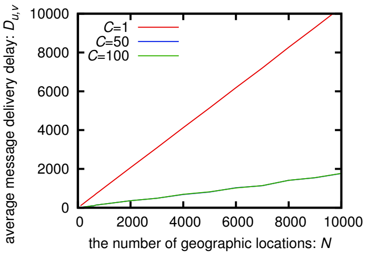

Fig. 1. Relation between the number N of geographic locations and average message delivery delay $D_{u,v}$.

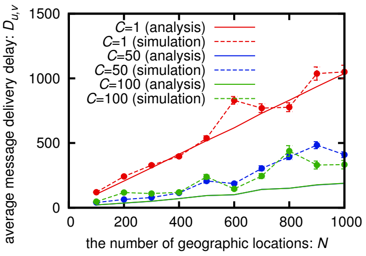

Fig. 2. Comparison of numerical results and simulation results.

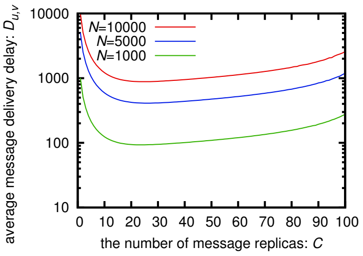

Fig. 3. Relation between the number C of message replicas and average message delivery delay $D_{u,v}$.

Figure 3 shows the average message delivery delay $D_{u,v}$ as a function of the number C of message replicas. In this figure, the load factor is set to $\alpha = 0.9$ to simulate highly loaded conditions. This figure clearly illustrates that the average message delivery delay $D_{u,v}$ is a concave function. The optimal number of message replicas is around $C = 20$ but it is almost independent of the network size N in this case.

Now we are ready to answer two research questions — how well or badly geographic DTN routing performs in a large-scale network and how the performance of geographic DTN routing is affected by the network topology.

Our observations regarding Fig. 1 answer the first question; somewhat surprisingly, geographic DTN routing is *scalable* in terms of the network size N since the average message delivery delay grows almost proportionally as the network size N increases, which is quite favorable property of geographic DTN routing.

Answering the second research question needs more explanation.

The second research question can be rephrased as follows: how the average message delivery delay $D_{u,v}$ in Fig. 1 is changed if *the average degree* is different (i.e., $\overline{k} \neq 6$) and if *the degree distribution* of graph G is not binomial (e.g., power-law distribution as in scale-free networks). Except highly loaded conditions, the dominant factor of the replica delivery delay X_i is the transfer delay X_i^T rather than the queueing delay X_i^Q in Eq. (4). As Eq. (12) implies, the transfer delay is mostly determined by the average hitting time $H_{u,v}$, which is then determined solely by the ratio of *the number* $|E|$ *of edges* to *the degree* $d(v)$ *of the destination node* v. Namely, in geographic DTN routing under random walk mobility, the network topology has *limited* impact on the performance of geographic DTN routing; the average message delivery delay is mostly determined by the degree of the destination node.

VII. CONCLUSION

In this paper, we have presented a hybrid model of multiple-copy geographic DTN routing, which is composed of a continuous $M/M/1/PS$ queueing model and discrete multiple random walks on a large-scale network. We have approximately derived the average message delivery delay in a geographic DTN routing under random walk mobility on a large-scale network. We have analytically obtained the optimal number of message replicas (i.e., the number of copies per a message). Our findings include that geographic DTN routing is scalable in terms of the network size N, and that the network topology has limited impact on the performance of geographic DTN routing except heavily loaded conditions.

Our future work includes extensive simulations under realistic scenarios to further validate our approximate analysis, and extension of our analysis to incorporate more realistic geographic DTN routing such as geometry-aware routing algorithms, and other mobility patterns than the random walk.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 16H02815.

REFERENCES

[1] M. J. Khabbaz, C. M. Assi, and W. F. Fawaz, “Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges,” *IEEE Communications Surveys and Tutorials*, vol. 14, pp. 607–640, Jan. 2012.

[2] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, “MaxProp: Routing for vehicle-based disruption-tolerant networks,” in *Proceedings of the 25th IEEE International Conference on Computer Communicaitions (INFOCOM 2006)*, pp. 1–11, Apr. 2006.

[3] S. C. Nelson, M. Bakht, and R. Kravets, “Encounter-based routing in DTNs,” in *Proceedings of the 28th IEEE International Conference on Computer Communications (INFOCOM 2009)*, pp. 846–854, Apr. 2009.

[4] J. Wu, M. Xiao, and L. Huang, “Homing spread: Community home-based multi-copy routing in mobile social networks,” in *Proceedings of the 32th IEEE International Conference on Computer Communications (INFOCOM 2013)*, pp. 2319–2327, Apr. 2013.

[5] M. Seki, K. Ogura, Y. Yamasaki, and H. Ohsaki, “Performance comparison of geographic DTN routing algorithms,” in *Proceedings of the IEEE Signature Conference on Computers, Software, and Applications (Student Research Symposium) (COMPSAC 2015)*, pp. 617–620, July 2015.

[6] D. Matsui, R. Hagihara, Y. Yamasaki, and H. Ohsaki, “Analysis of geographic DTN routing under random walk mobility model,” in *Proceedings of the 41th IEEE Signature Conference on Computers, Software, and Applications (COMPSAC 2017)*, pp. 538–547, July 2017.

[7] R. Monteiro, W. Viriyasitavat, S. Sargento, and O. K. Tonguz, “A graph structure approach to improving message dissemination in vehicular networks,” *Wireless Networks*, vol. 23, pp. 2145–2163, Apr. 2017.

[8] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports*, vol. 424, no. 4, pp. 175–308.

[9] C. Stegehuis, R. van der Hofstad, and J. S. H. van Leeuwaarden, “Epidemic spreading on complex networks with community structures,” *Scientific Reports*, vol. 6, July 2016.

[10] M. Shahzamal, M. F. Pervez, M. A. U. Zaman, and M. D. Hossain, “Mobility models for delay tolerant network: A survey,” *International Journal of Wireless & Mobile Networks (IJWMN)*, vol. 6, pp. 121–134, Aug. 2014.

[11] L. Lovász, “Random walks on graphs: a survey,” *Combinatorics, Paul Erdos is eighty, Keszthely*, vol. 2, pp. 1–46, 1993.

[12] U. von Luxburg, A. Radl, and M. Hein, “Hitting and commute times in large graphs are often misleading.” https://arxiv.org/abs/1003.1266.pdf, May 2011.

[13] P. Erdös and A. Rényi, “On random graphs I,” *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297, 1959.