Adaptive Fractal-like Network Structure for Efficient Search of Targets at Unknown Positions and for Cooperative Routing

Yukio Hayashi and Takayuki Komaki,
Japan Advanced Institute of Science and Technology
Email: yhayashi@jaist.ac.jp, takayuki.k@mocha.ocn.ne.jp

***Abstract*—From viewpoints of complex network science and biological foraging for communication networks, we propose a system model of scalable self-organized geographical networks, in which the proper positions of nodes and the network topology are simultaneously determined according to population. The fractal-like network structure is constructed by iterative divisions of rectangles for load balancing across nodes, in order to adapt to territory changes. In numerical simulations, we show that, for searching targets concentrated around high population areas, the naturally embedded fractal-like structure by population has higher efficiency than the conventionally optimal strategy on a square lattice. The adaptation of network structure to the spatial distribution of realistic communication requests gives such a high performance.**

Keywords -self-organized design, spatially inhomogeneous communication requests, random walk, Lévy flight, cooperative message ferries.

I. INTRODUCTION

The size and complexity of communication and transportation networks are growing year by year for the increase of users, communication requests, mobility, and technological innovations in real-world sensing or handling of rich contents. For the design and control of growing networks, scalability, adaptivity, and self-organization will be more required in the near future. Here, we consider the adaptive network structure [1] to be suitable for searching a target on the network embedded in a space with population density, because searching (or routing) is one of the important basic tasks to establish a connected path on a network.

Many network infrastructures: power grids, airline networks, and the Internet, are embedded in a metric space, and long-range links are relatively restricted [2], [3] for economical reasons. The spatial distribution of nodes is neither uniformly at random nor on a regular lattice, which is often assumed in the conventional network models. In real data, a population density is mapped to the number of router nodes on Earth [2]. Similar spatially inhomogeneous distributions of nodes are found in air transportation networks [4] and in mobile communication networks [5]. Thus, it is not trivial how to locate nodes within a space using patterns of points. Point processes in spatial statistics [6] provide models for irregular patterns of points in urban planing, astronomy, forestry, or ecology, such as spatial distributions of rainfall, germinations, plants, and animals. The processes assume homogeneous Poisson and Gibbs distributions to generate a pattern of random packing or independent clustering, and to estimate parameters of competitive potential functions in a territory model for a given statistical data, respectively. However, rather than random pattern and statistical estimation, we focus on a self-organized network infrastructure by taking into account realistic spatial distributions of nodes and communication requests. In particular, we aim at developing adaptive and scalable networks by adding the links between proximity nodes according to the increase of communication requests. Because a spatial distribution of communication requests affects the proper positions of nodes, which control both the load of requests assigned to each node (e.g., assigned at the nearest access point of node as a base-station from a user) and the communication efficiency depending on the selection of routing paths.

Thus, we propose a scalable self-organized geographical network, considering an interrelation of routing algorithms in computer science, biological foraging, and complex network science. Complex network science is a groundbreaking science that has emerged from a physical society about ten years ago for understanding the common network structures in social, technological, and biological systems [7], [8], [9], and the fundamental generation mechanisms. We show that the naturally embedded fractal-like structure in the proposed network [10] is suitable for searching inhomogeneously distributed targets more efficiently than the square lattice tracked by the Lévy flights, which is known as an optimal biological search for homogeneously distributed targets [11]. Moreover, we investigate the performance for a cooperative routing method in the fractal-like network, as an extension of the conference paper [1]. We emphasize that a spatially inhomogeneous distribution of communication requests is important [10] for a realistic situation according to population density, and that an adaptive network structure to population is self-organized.

The organization of this paper is as follows. For starting a discussion, in Section II, we mention the miminimum

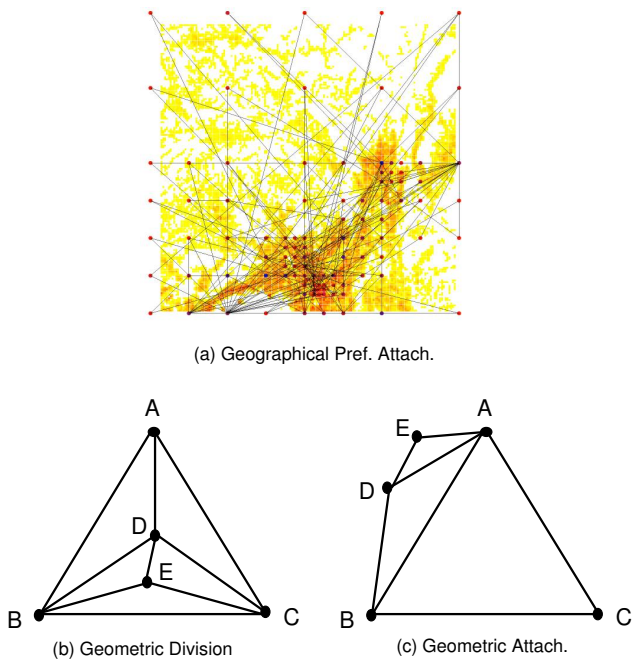
necessity of related works. Because we treat several different topics and the interrelations of network self-organization, routing methods by a determinant or stochastic walker, biological search strategy, and cooperative routing by multiple walkers. To avoid confusion, each detail is explained in the most related section. In Section III, we briefly review the conventional geographical network models, and propose a new model based on iterative divisions of rectangles. In Section IV, we show the basic search performance of a random walk in the fractal-like network structure for targets at unknown positions. In Section V, we investigate the routing property by using cooperative agents on the network structure, comparing that with the Lévy flights on a square lattice. In Section VI, we summarize these results and mention further studies.

II. Related Work For Routing and Searching

For routing in ad hoc networks, global information, e.g., a routing table in the Internet, cannot be applied, because many nodes and connections between them are likely to change over time. Although there are many protocols [12], [13] for energy saving, mobile networks, GIS-based location awareness, QoS, and wireless sensor devices, we restrict strongly related ones to our discussion.

In early works on computer science, some decentralized routing methods were developed to reduce energy consumption in sensor or mobile networks. However they lead to the failure of guaranteed delivery [14]; in the flooding algorithm, multiple redundant copies of a message are sent and cause congestion, while greedy and compass routings may occasionally fall into infinite loops or into a dead end. We do not need to be particular about these simple and energy saving methods in the current and future technologies. At least, it is better to guarantee the delivery. In complex network science, other no-failure efficient decentralized routing methods have been also proposed. The stochastic methods by using local information of the node degrees and other measures are called preferential [15] and congestion-aware [16], [17] routings as extensions of a uniformly random walk.

Decentralized routing has a potential performance to search a target whose position is unknown in advance. Since this situation looks like foraging, the biological strategy may be useful for the efficient search. We are interested in a relation between the search and the routing on a spatially inhomogeneous network structure according to population. Many experimental observations for biological foraging found the evidence in favor of anomalous diffusion in the movement of insects, fishes, birds, mammals and human being [11]. As a consistent result, it has been theoretically analyzed for a continuous space model that an inverse square distribution of flight lengths is an optimal strategy to search sparsely and randomly located targets on a homogeneous space [18]. The discrete space models on a regular lattice [19] and the defective one [20] are also discussed. Such behavior is called *Lévy flight* characterized by a distribution function $P(l_{ij}) \sim l_{ij}^{-\mu}$ with $1 < \mu \leq 3$, where l_{ij} is a flight length between nodes i and j in the stochastic movement for any direction. The values of $\mu \geq 3$ lead to Brownian motions, while $\mu \to 1$ to ballistic motions. The optimal case is $\mu \approx 2$ for maximizing the efficiency of search. Here, we assume that the mobility of a node is ignored due to a sufficiently slow speed in comparison with the communication process. In current or future technologies, wide-area wireless connections by directional beams will be possible, the modeling of unit disk graph with a constant transmission range is not necessary. Thus, as a system model, we consider efficient search and routing on an adaptive fractal-like network structure to spatially inhomogeneous communication requests.



(a) Geographical Pref. Attach.

(b) Geometric Division

(c) Geometric Attach.

Figure 1. Typical construction methods of geographical networks. At each time step, a new node is added at a random position. Then, (a) the new node i is linked to an existing node j chosen with a probability $\Pi_j \propto d_{ij}^{-\alpha} pop_j^{\beta} k_j^{\gamma}$, where α, β, and γ are real parameters, d_{ij} denotes the distance between nodes i and j, pop_j denotes the population in the node j's territory, and k_j is its degree. The territory is a merged area of mesh blocks, which nearest access node is j. The gradation (from while, yellow, orange, to red) of background mesh blocks on a $L \times L$ square lattice is proportional to the value of population given from a census data. The case of $\alpha = \beta = 0$ and $\gamma > 0$ is the degree based model, and the case of $\alpha, \gamma > 0$ and $\beta = 0$ is a combination of degree and distance based model. On the other hand, the new node (D or E) is set at random (b) in a chosen triangle or (c) outside of a chosen edge, and linked to (b) the three nodes of the randomly chosen triangle or (c) both ends of the randomly chosen edge. The initial configuration consists of triangles.

III. Geographical Networks

We introduce geographical network models proposed in complex network science, which aims to elucidate a fun-

damental mechanism for generating an efficient network structure in a distributed manner.

A. Conventional Models

Geographical constructions of complex networks have been proposed so far. Figure 1 (a) and (b)(c) shows the typical methods. It is well known that the preferential attachment [7] is fundamental to construct a scale-free (SF) network that follows a power law degree distribution found in many real systems [21], [22]. As a generation mechanism of geographical SF networks, a spatially preferential attachment is applied in some extensions [23], [24], [25], [26], [27] from the topological degree based model [28] to a combination of degree and distance based model (Figure 1(a)). However, this construction tends to have many long links, which are wasteful. The original degree based preferential attachment is known as "rich gets richer" rule that means a higher degree node tends to get more links. It is a surprising thing that inhomogeneous complex network structures emerge from such a simple rule. On the other hand, geometric construction methods have also been proposed (Figure 1(b)(c)). They have both small-world [29] and SF structures generated by a recursive growing rule for the division of a chosen triangle [30], [31], [32], [33] or for the attachment aiming at a chosen edge [34], [35], [36] in random or hierarchical selections. Here, small-world means that the average path length counted by hops between two nodes is very small as $O(\log N)$ even in a large size N defined by the total number of nodes. These geometric models are proper for the analysis of degree distribution due to the regularly recursive generation process. Although the position of a newly added node is basically free as far as the geometric operations are possible, it has no relation to population. Considering the effects of population in a geographical network is necessary to self-organize a spatial distribution of nodes that is suitable for socioeconomic communication and transportation requests. Moreover, in these geometric methods, narrow triangles with long links tend to be constructed, and adding only one node per step may lead to exclude other topologies from the SF structure. Unfortunately, SF networks are extremely vulnerable against the intentional hub attacks [37]. We should develop other models of self-organized networks distinct from the conventional models; e.g., a better network without long links can be constructed by subdivisions of equilateral triangles, which is a well balanced (neither fat nor thin) shape for any directions as shown in Figure 2(a). In the network without long links, a node with a small degree does not become hub, therefore the attack vulnerability of connectivity disappears.

B. Generalized MSQ Network

Thus, we have considered the multi-scale quartered (MSQ) network model [38], [39]. It is based on a stochastic construction by a self-similar tiling of primitive shape.

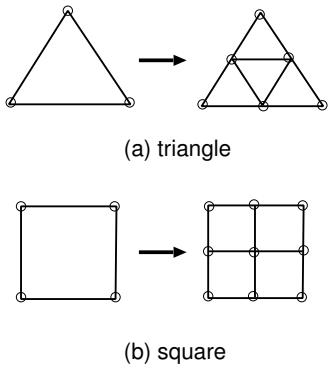


Figure 2. Basic process of the division.

Figure 2(a)(b) shows the basic process of division in the tiling of equilateral triangle or square. At each time step, a face is chosen proportionally to the population in the space. Then, the chosen face is divided into four smaller equilateral triangles or squares. This process is repeated. The MSQ networks without hub nodes have several advantages such as the strong robustness of connectivity (due to the small degrees) against node removals by random failures and intentional attacks, the bounded short path as $t = 2$-spanner [40], and the efficient face routing by using only local information. The t-spanner means that the length of shortest distance path (defined by the sum of link lengths on the path) between nodes u and v is bounded at most t times the Euclidean distance d_{uv} of the corresponding straight line between them. In the face routing, the shortest distance path can be found on the edges of faces that intersect the straight line, since the MSQ network is planner, which is also suitable for avoiding the interference among wireless beams. Furthermore, the MSQ networks are more efficient (economic) with shorter link lengths and more suitable (tolerant) with lower load for avoiding traffic congestion [39] than the state-of-the-art geometric growing networks [30], [31], [32], [33], [34], [35], [36] and the spatially preferential attachment models [23], [24], [25], [26], [27] with various topologies ranging from river to SF geographical networks. However, in the MSQ networks, the position of a new node is restricted on the half-point of an edge of the chosen face, and the link length is proportional to $(\frac{1}{2})^H$ where H is the depth number of iterative divisions. Thus, from square to rectangle, we generalize the division procedure as follows. Figure 3 illustrates it.

- Step 0: Set an initial square, in which the candidates of division axes are the segments of an $L \times L$ lattice (Figure 3(a)).
- Step 1: At each time step, a face is chosen with a probability proportional to the population counted in the face covered by mesh blocks of a census data (Figure 3(b)).
- Step 2: Four smaller rectangles are created from the

division of the chosen rectangle by horizontal and vertical axes. For the division, two axes are chosen by that their cross point is the nearest to the population barycenter of the face (Figure 3(c)).

Step 3: Return to Step 1, while the network size N does not exceed a given size.

Note that the maximum size N_{max} depends on the value of L; the iteration of division is finitely stopped, since the extreme rectangle can not be divided any longer when one of the edge lengths of rectangle is the initial lattice's unit length. We use the population data on a map in $80km^2$ of 160×160 mesh blocks ($L = 160$) provided by the Japan Statistical Association. Of course, other data is possible.

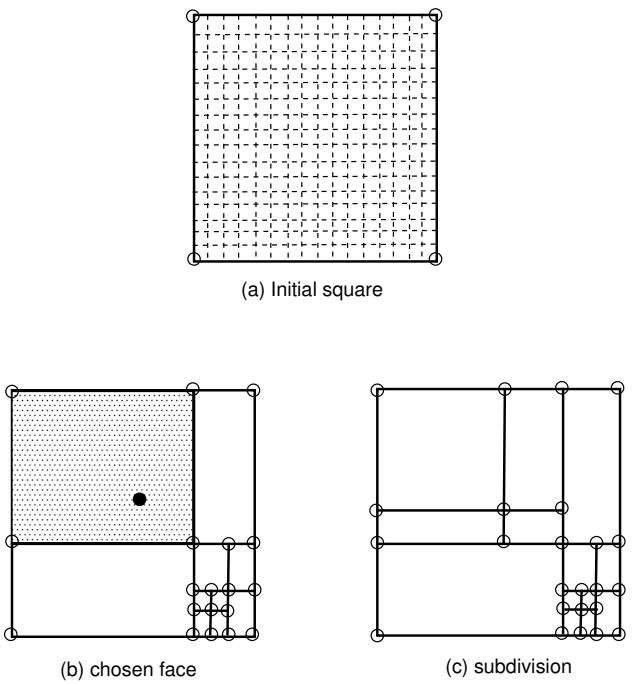


Figure 3. Division procedure in generating a generalized MSQ network. (a) Initial configuration: the outer square of 4 nodes and 4 links. For division, dashed-lines represent the segments defined by the edges of $L \times L$ mesh blocks. To each mesh block whose right-bottom corner is at $1 \leq x \leq L$ and $1 \leq y \leq L$, a value of population $pop_{x,y}$ is assigned by a census data. (b) As an example, a (shaded) face f is chosen with a probability $\propto \sum_{x,y \in A_f} pop_{x,y}$ at the 4th time step. Where A_f denotes the set of x-y coordinate values included in the face f. (c) Then, the horizontal and vertical axes, which cross point is nearest to the population varycenter (black filled circle) of face f, are selected, and divide the chosen face into four smaller ones.

It is worth noting that the positions of nodes and the network topology are simultaneously determined by the divisions of faces within the fractal-like structure. There exists a mixture of sparse and dense parts of nodes with small and large faces. Moreover, while the network is growing, the divisions of faces perform a load balancing of nodes in their adaptively changed territories for the increase of population. Such a network is constructed according to a spatially inhomogeneous distribution of population, which is proportional to communication requests in a realistic space. In the following, we show the naturally embedded fractal-like structure is suitable for searching targets. Moreover, we apply the good property to a routing task in Section V.

IV. Basic Search Performance

As a preliminary, we consider the preferential routing [15] which is also called α-random walk [41]; The forwarding node j is chosen proportionally to K_j^α by a walker in the connected one hop neighbors $\mathcal{N}_i$ of its resident node i of a walker (packet), where K_j denotes the degree of node j and α is a real parameter. We assume that the start position of walker is set to the nearest node to the population barycenter of the initial square. Figure 4 shows the length distribution of visited links. The dashed lines in log-log plot suggest a power law, for which the exponents estimated as the slopes by a mean-square-error method are 2.336, 2.315, and 2.296 for $\alpha = 1, 0, -1$, respectively. These values are close to the optimal exponent $\mu \approx 2$ [18], [19] in the Lévy flight on a square lattice. The exponents for the α-random walks slightly increase as the network size N becomes larger. Here, the case of $\alpha = 0$ shows the length distribution of existing links in a network. Since the stationary probability of incoming at node j is $P_j^\infty \propto K_j^{1+\alpha}$ [42], especially at $\alpha = 0$, each of the connected links to j is chosen at random by the probability $1/K_j$ for the leaving from j, therefore a walker visits each link at the same number. Figure 5 shows that the frequency of visited links by the α-random walks at $\alpha = \pm 1$ is different even for the degrees 3 and 4 in a generalized MSQ network. On the thick lines, a walker tends to visit high population (diagonal) areas colored by orange and red in the case of $\alpha = 1$, while it tends to visit low population peripheral (corner) areas in the case of $\alpha = -1$. Thus, the case of $\alpha = 1$ is expected to selectively cover high population areas, which has a lot of communication requests in cities. Note that the absolute value of α should be not too large, since a walker is trapped a long time between high/low degree nodes as the phenomena does not contribute to the search of targets.

We investigate the search efficiency for the α-random walk on a generalized MSQ network, and compare the efficiency with that for Lévy flights on a $L \times L$ square lattice with periodic boundary conditions [19]. As shown in Figure 6, a walker constantly looks for targets (destination nodes of packets) scanning on a link between two nodes in the generalized MSQ network. If a target exists in the vision area of r_v for the up/down/left/right directions from the center position, a walker gets it and return to the position on the link for continuing the search on the same direction. When more than one target exist in the area, a walker gets all of them successively in each direction, and return to the position. Only at a node of rectangle, the search direction is changeable along one of the connected links.

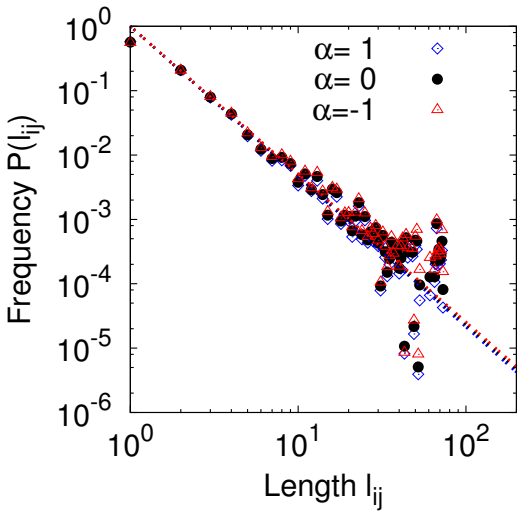


Figure 4. Length distribution of visited links on generalized MSQ networks by an α-random walker in 10^6 time steps. The marks of blue diamond, black circle, and red triangle correspond to the cases of $\alpha = 1, 0, -1$, respectively. These results are obtained by the average of 100 networks for $N = 2000$.

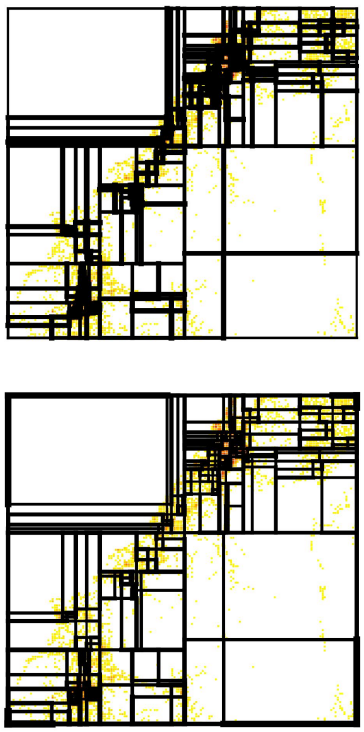


Figure 5. Visualization examples of the visited links by α-random walks at $\alpha = 1$ (Top) and $\alpha = -1$ (Bottom) on a generalized MSQ network for $N = 500$. The thickness of link indicates the frequency of visiting in 10^6 time steps. From light to dark: white, yellow, and orange to red, the color gradation on a mesh block is proportionally assigned to the population data. Many nodes represented as cross points of links concentrate on high population (dark: orange and red) areas on the diagonal direction. In the upper left and lower right of square, corner triangle areas lighted by almost white are the sea of Japan and the Hakusan mountain range.

Thus, the search is restricted on the edges of rectangle in the generalized MSQ network. While the search direction of a Lévy flight on the square lattice [19] is selectable from four directions of horizontal and vertical at all times after getting a target in the scanning with the vision area of r_v, moreover, the length of scan follows $P(l_{ij}) \sim l_{ij}^{-\mu}$, $l_{ij} > r_v$. We set a target at the position chosen proportionally to the population around a cross point in $(L+1)^2$, for which the population is defined by the average of four values in its contact mesh regions. In particular, we discuss the destructive case [19]: once a target is detected by a walker, then it is removed and a new target is created at a different position chosen with the above probability. Similar results to below in this section are obtained for the non-destructive case [10].

The search efficiency [18], [19], [20] is defined by

$$\eta \stackrel{\text{def}}{=} \frac{1}{M} \sum_{m=1}^{M} \frac{N_s}{L_m}, \tag{1}$$

$$\lambda \stackrel{\text{def}}{=} \frac{(L+1)^2}{N_t 2 r_v}, \tag{2}$$

where L_m denotes the traversed distance counted by the lattice's unit length until detecting $N_s = 50$ targets from the total N_t targets in the mth run. We consider a variety of $N_t = 60, 100, 200, 300, 400$, and 500 for investigating the dependency of the search efficiency on the number N_t of targets. The quantity λ represents the mean interval between two targets for the scaling of efficiency by target density. We set $M = 10^3$ and $r_v = 1$ for the convenience of simulation. Intuitively, the sparse and dense structures according to the network size N have the advantage and disadvantage in order to raise the search efficiency in the generalized MSQ network. Although the scanned areas are limited by some large rectangle holes as N is small, a walker preferably visits the high population areas that include many targets. While the scanned areas are densely covered as N is large, the search direction is constrained on long links of a collapse rectangle, therefore it is rather hard for a walker to escape from a local area in which targets are a few.

We compare the search efficiency of α-random walks on the generalized MSQ networks with that of the Lévy flights on the square lattice. Figure 7 shows typical trajectories until detecting $N_s = 50$ targets. On the generalized MSQ network and the square lattice, a walker tends to cover a local area with high population and a wider area, respectively. Without wandering in peripheral wasteful areas, the generalized MSQ network has a more efficient structure than the square lattice for detecting many targets concentrated on the diagonal areas. Here the exponent $\mu = 1.8$ of Lévy flight corresponds to the slope of $P(l_{ij})$ in the generalized MSQ network at the

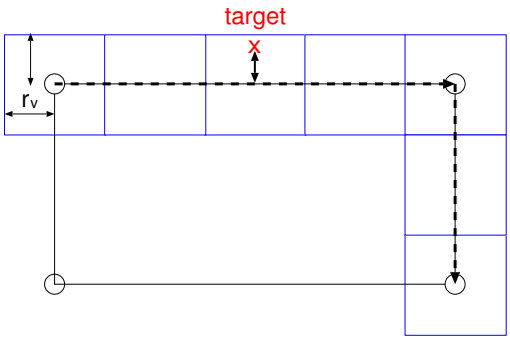
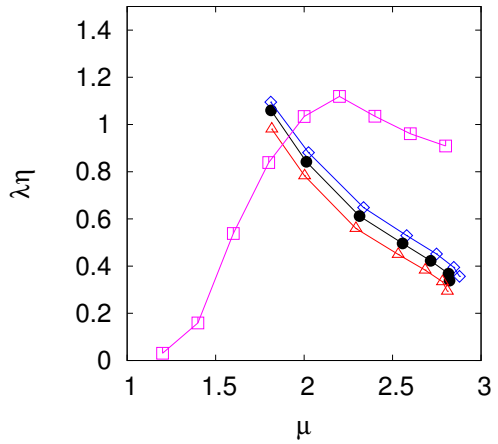


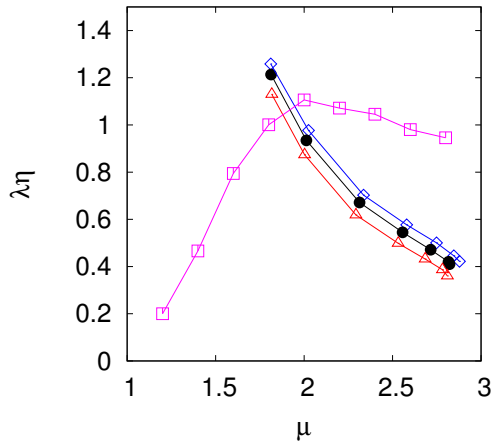
Figure 6. Searching in a generalized MSQ network. Each blue square represents a vision area, and is scanned (from left to right, from top to bottom in this example) by the walker on an edge between two nodes (denoted by circles) of a rectangle. For a target in the area, the walker moves to get it and returns on the link.

Figure 7. Trajectories of a random walk (Top) at $\alpha = 0$ on a generalized MSQ network for $N = 500$ and of a Lévy flight (Bottom) for $\mu = 1.8$ on the square lattice with periodic boundary conditions until detecting $N_s = 50$ targets in $N_t = 200$. Black circle, red circles, and gray rectangle marks denote the start point at the population barycenter, the existing targets, and the removed targets after the detections, respectively. Note that a walker can travel back and forth on a link in the connected path.

optimal size $N = 500$ for the search efficiency. As shown in Figure 8(a)(b), the generalized MSQ networks of $N = 500$ (the diamond, circle, and triangle marks are sticking out at the left) have higher efficiency than the square lattice (the rectangle mark). For the cases with many nodes of $N \geq 1000$, the efficiency decreases more rapidly than that of the Lévy flight, however this phenomenon means that an extremely large network size is wasteful and unnecessary to get a high search performance in generalized MSQ networks. When the number N_t of targets increases in cases from Figure 8(a) to (b), the curves are shifted up, especially for the generalized MSQ networks. The peak value for $N_t = 200$ is larger than the optimal case of the Lévy flight at $\mu = 2.0$. Therefore denser targets to that extent around $N_t = 200$ is suitable, although a case of larger $N_t > 300$ brings down the search efficiency even for inhomogeneously distributed targets.



(a) $N_t = 100$



(b) $N_t = 200$

Figure 8. The scaled efficiency $\lambda\eta$ vs. the exponent μ. The marks of blue diamond, black circle, and red triangle correspond to the cases of $\alpha = 1, 0, -1$, respectively, in which the increasing values of μ are estimated for generalized MSQ networks at $N = 500, 1000, 2000, 3000, 4000, 5000,$ and 5649: N_{max} from left to right. The magenta rectangle corresponds to the case of Lévy flights on the square lattice. These results are obtained by the average of 100 networks.

In more details, Figure 9 shows the effect of the number N_t of targets on the search efficiency $\lambda\eta$. The efficiency firstly increases, then reaches at a peak, and finally decreases

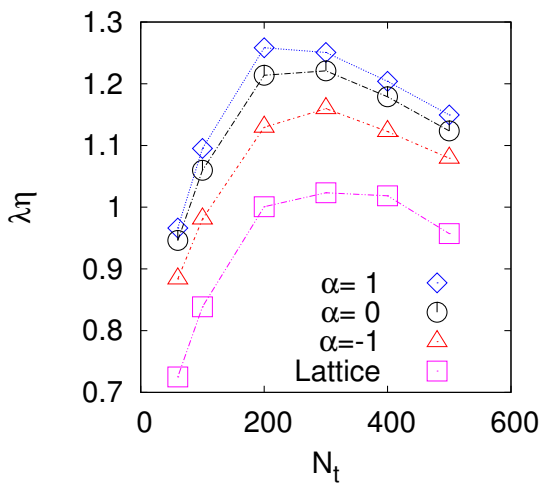


Figure 9. The number N_t of targets vs. the scaled efficiency $\lambda\eta$ of α-random walks on the generalized MSQ networks for $N = 500$ and of the corresponding Lévy flights for $\mu = 1.8$ (see Figure 8) on the square lattice. The maximum (optimal) efficiency appears in $N_t = 200 \sim 300$. These results are obtained by the average of 100 networks.

for setting more targets. This up-down phenomenon is caused by a trade-off between L_m and N_t in Eqs. (1) and (2). Please note that the case of size $N < 500$ is omitted for the generalized MSQ networks. Because sometimes the process for detecting targets until N_s is not completed, moreover, the variety of link lengths is too little to estimate the exponent as a slope of $P(l_{ij})$ in the log-log scale. In other words, the estimation is inaccurate because of the short linear part.

V. Routing by Message Ferries

When a communication network is often disconnected but resilient due to node mobility, limited radio power, node or link failure, etc., it is known as a Delay/Disruption Tolerant Network (DTN) where a mobile device or software agent temporary stores and carries local information for forwarding messages until an end-to-end route is re-established or re-generated. It is used in disasters, battlefields, and vehicular communications. There are many protocols in the concept of DTN routings [43], [44]. A message ferrying scheme is one of the DTN routing strategies, in which a device or agent called ferry stores, carries, and forwards messages in partitioned ad hoc networks. It is classified into a single ferry [45] or multiple ferries, stationary or mobile node, node-initiated or ferry-initiated moving to communicate [46], single-route or multi-routes, and node relaying or ferry relaying [47] according to the protocol components: a movement of ferry, interactions between node and ferry or between ferries, how much and which type of local information is stored in a ferry or at a node, and so on.

We focus on cooperative multiple ferries as software agents, because the ferries interact asynchronously through a mediator node sharing partial information or exchanging it for their routing tasks. This method can avoid the problem of very rare encounter between ferries because of their random walks. During a routing, we assume a network is fixed to distinguish the effects by the network structure and by the disconnections on the performance, since we consider the network structure is a primal factor to control a ferry's movement. In addition, we distinguish between how to determine a route and how to deliver a message (data packets), thus we do not care whether or not a ferry should move with its message. The appropriative delivery depends on the ability of devices, the amount of message, and communication environment. However, our approach will be applicable to an opportunity networking with node mobility. Note that a Lévy walk of a single ferry is applied for searching on a continuous space with the Euclidean distance in order to maximize the opportunity of encounter with the destinations of mobile nodes [48], though the problem setting is different from ours. In the following, we consider only the case of $\alpha = 0$ in the α-random walks, to simplify the discussion, since the difference for the cases of $\alpha = \pm 1$ is small.

A. Multiple Message Ferries Routing

We explain the outline of routing algorithm. Note that a ferry has no vision area in the routing problem unlike the searching problem discussed in the previous section. In addition, a walker moves to get many targets in the searching problem, while multiple ferries move to cooperatively find paths between source and destination nodes in the routing problem.

Initially, there is no label at each node. Communication requests by different pairs of source s and destination d are labeled at a node. In other words, as a mediator between ferries, each node handles more than one requests that are carried from ferries. Similarly, a ferry can carry more than one requests. Without global information, a node visited by a ferry A at time t memorizes the node $n_A(t-1)$ for each ferry in the connected neighbors, where $n_A(t-1)$ denotes the visited node by A at $t-1$. While a ferry memorizes a set of passed links as the history in a limited size of buffer. Thus, using only partial information, a path between any two nodes is found as follows.

1) **Comm. Request:** A pair of s and d nodes is chosen proportionally to the population counted in the territory of each node (defined by the nearest access point) for a census data referred in Subsection III. The generation rate R is the number of generated s-d pairs in the network per time step.
   After the generation at node s, the communication request $REQ(s, d)$ is put on hold until a ferry encounters it for the carrying, because each node is fixed.

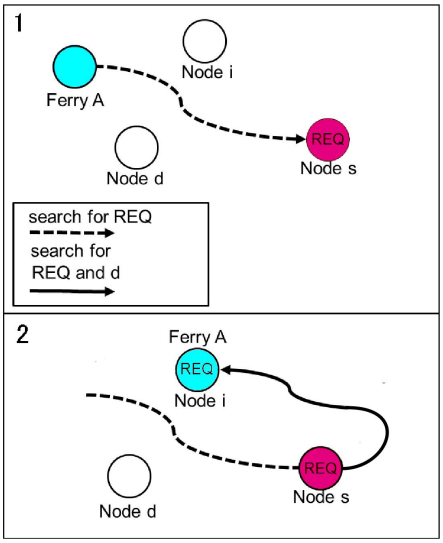


Figure 10. Ferry's modes. (Top) In free-mode, a ferry A encounters the source node s for a request: $REQ(s,d)$ on its walk. (Bottom) In search-mode after the encounter, the ferry A carries the request $REQ(s,d)$ to other nodes, e.g., i, for searching the destination node d.

2) **Ferry's Action:** At each time step, each ferry walks at random on a generalized MSQ network, storing the passed links into its stack-based buffer. Using a set of the stored links, it interacts with the visited node as mentioned in 3).
A ferry has two modes: free and search as shown in Figure 10.
   - Free-Mode: This mode is preliminary to searching, but may contribute to a cooperation between ferries: please see 4). When a ferry of free-mode encounters a source node s or requests handled and labeled at a visiting node, the mode is changed to search-mode.
   - Search-Mode: A ferry of search-mode carries a set $\{REQ(s,d)\}$ of requests to a node i through the random walk.
     Moreover the ferry A asks whether or not the node i knows the $REQ(s,d)$ in its label. If the answer is "NO," the visiting node i is labeled by the $REQ_A(s,d)$. Here, the suffix $_A$ is added to distinguish which ferry carries the $REQ(s,d)$ for backtracking in the path finding. This inquiry is tested for all requests carried by the ferry.
     Of course, if the visiting node is d, then a path between s and d is found: please see 5).

3) **Node Mediator:** A ferry interacts with the visiting node at a time. For each request $REQ(s',d')$ handled at the node, the node asks whether or not the ferry has a link to d' in the buffer stored as the visiting history. If the answer is "YES," go to 4).
   Some requests which the ferry does not have are copied from the node to the ferry for the carrying.

4) **Cooperation:** As shown in Figure 11, when a ferry B, which has a link to d in its buffer, visits a node i labeled by $REQ_A(s,d)$, a path between s and d is found. Because the existence of label $REQ_A(s,d)$ means that the node i is already visited by another ferry A, which comes from s (In more detail, a path from i to s is obtained from the node's information through switching ferries A, C, ..., Z, which carry the $REQ(s,d)$ via intermediate nodes from s to i', from i' to i'', ..., and to i.).

5) **Path Finding:** A path between s and d is found in a subgraph, which consists of the links (including a link to the destination d) in the ferry's buffer and the backward connections of $\{n_A(t-1)\}$ nodes for the ferry A until reaching the source s.
   If a ferry A starting from s visited d, then the ferry's buffer is unnecessary for the path finding. In other words, this case has no cooperation, or is equivalent to the case of zero buffer size.

After the find of a path between s and d, the multiple labels of $REQ_A(s,d), REQ_B(s,d), \ldots$ by ferries $A, B, \ldots$ at a node are deleted in a distributed manner, if the forward connections of $n_A(t+1), n_B(t+1), \ldots$ nodes are memorized for each related ferry $A, B, \ldots$ to the request $REQ(s,d)$.

We investigate the average time step and movement distance until a ferry encounters a source node s from the generation of $REQ(s,d)$. On a movement, the distance means the sum of link lengths or flight lengths counted by the unit of the square lattice. Figure 12(a)(b) shows that both the time and the distance decrease as the number m of ferries increase. The slope near $1/m$ is consistent with the effect of speed-up in parallel walks [49]. Figure 13(a)(b) shows the average time and distance for a path finding. They roughly follow the $1/m$ property, but their slopes depend on the buffer sizes. As the buffer ratio becomes large, both the time and the distance decrease by the effect of cooperation between ferries. These results do not depend on the size N and the packet generation ratio R.

Table I
RELATION OF THE NETWORK SIZE N AND THE BUFFER SIZE FOR THE RATIO BR:2.0. NOTE THAT THE TOTAL NUMBER OF LINKS IS PROPORTIONAL TO N IN A GENERALIZED MSQ NETWORK.

N	buffer size
500	165
1000	333
2000	673
3000	1019

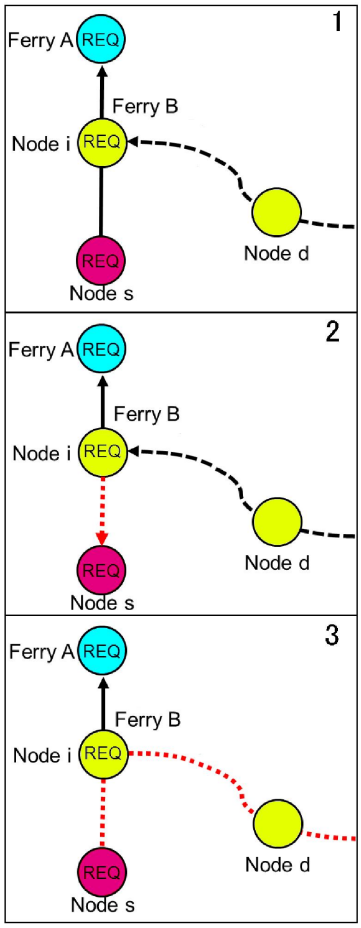
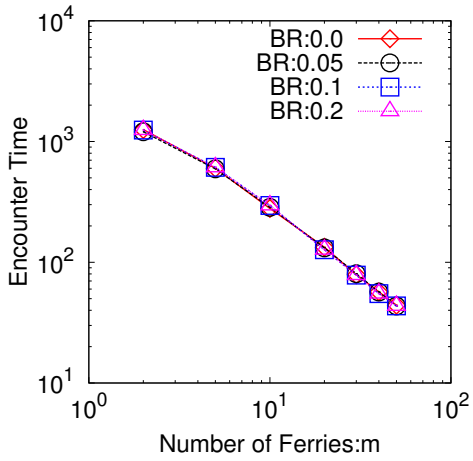
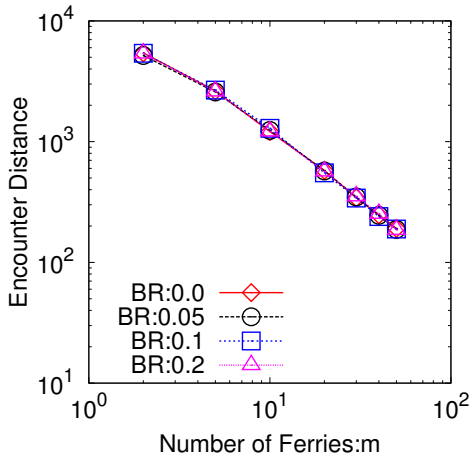


Figure 11. Cooperation of two ferries. When the ferry B visits a node i, which was visited by the ferry A with the request $REQ(s,d)$, a path between s and d is found from the concatenation of A's information and B's information in the situations: (Top) 1: The ferry B notifies the visiting experience at d to the node i, (Middel) 2: It backtracks the links from i to s on the red line by using $\{n_A(t-1)\}$ in the token relay via nodes, (Bottom) 3: From the subgraph that consists of the above gathered links and the B's link set, a path between s and d is calculated, e.g, by a criterion of the shortest distance. In this case, the ferry B does not have the $REQ(s,d)$, however it already visited d and stored the set of links connected to d into its buffer.



(a) Time



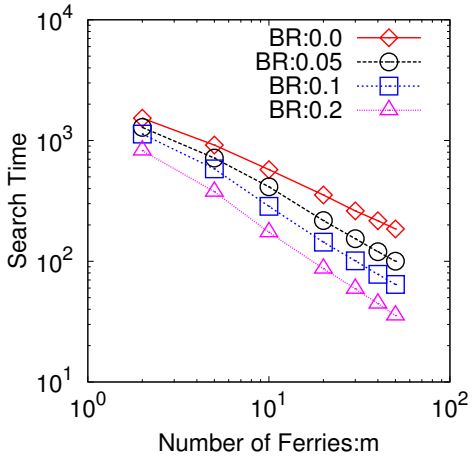
(b) Distance

Figure 12. Average time and distance over 50 realizations until a ferry of free-mode encounters a node s at which a request $REQ(s,d)$ is generated. BR denotes the buffer ratio as the maximum stored size of links in a ferry to the total number of links in the network. Although the encounter time does not depend on the values of BR, because a cooperation of ferries does not start, they are marked to be compared with the subsequent result in Figure 13. Here, the generation rate of requests is $R = 0.01$, and the network size is $N = 1000$.

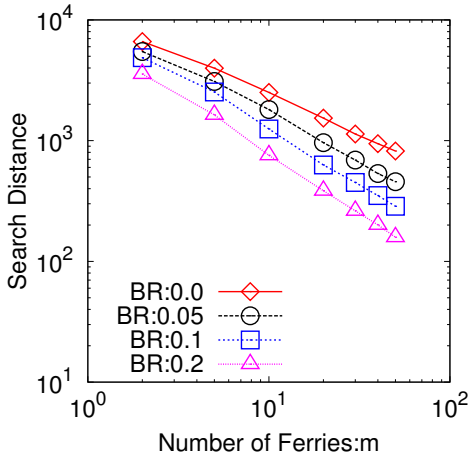
B. Comparison with Lévy Flights

In the cooperative message ferries scheme, we compare the performance of routing by random walks on a generalized MSQ network with that by Lévy flights on a $L \times L$ square lattice with periodic boundary conditions. The lattice is a background virtual space to determine a ferry's movement according to the Lévy flight. Note that only part 2) **Ferry' Action** is replaced in the routing algorithm for the Lévy flight version.

We investigate the average behavior over 50 realizations for each case of $m = 20$ ferries (in order to save computation time) in the combinations of the exponent μ of Lévy flights or the corresponding size of generalized MSQ networks and the buffer ratio (BR). The following simulation conditions are applied in both cases of random walks on a generalized MSQ network and Lévy flights on the lattice. For the generation of a communication request with rate $R = 0.01$, a s or d node is not able to set all lattice points but restricted on the node of a generalized MSQ network, and chosen proportionally to the population in the territory of each node in the network. Because a ferry that walks on the network can not visit any lattice point, in contrast, a ferry that moves according to the Lévy flight can visit any node in the network. The BR is set as 0.2 based on the results in Figures 12 and 13. Note that a larger buffer size tends to be effective in the routing in both short time and distance, however it

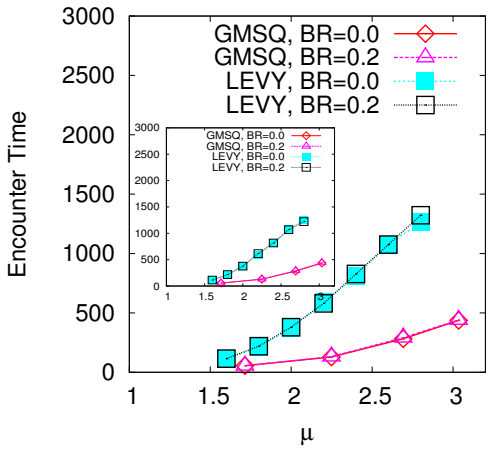


(a) Time

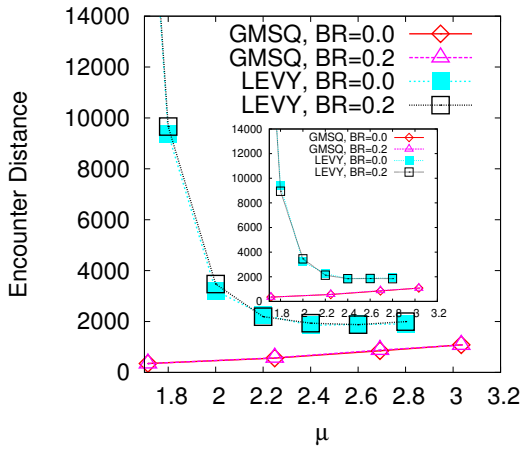


(b) Distance

Figure 13. Average time and distance over 50 realizations until a ferry finds a path between s and d from the encounter with a request $REQ(s,d)$. BR denotes the buffer ratio as the maximum stored size of links in a ferry to the total number of links in the network. Here, the generation rate of requests is $R = 0.01$, and the network size is $N = 1000$.



(a) Time

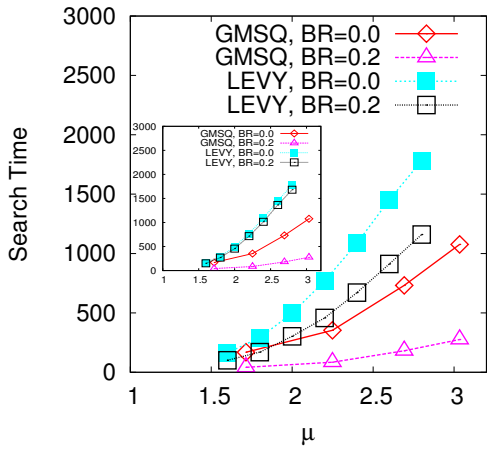


(b) Distance

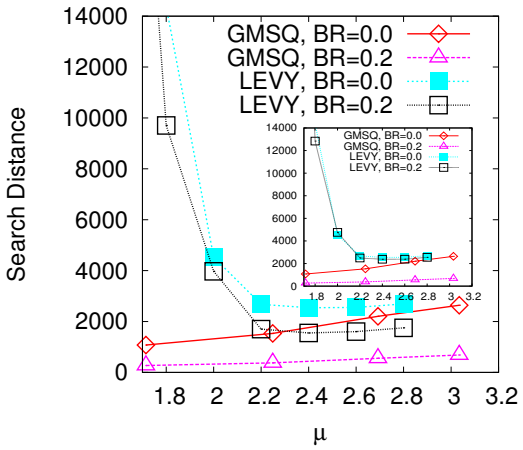
Figure 14. Average encounter time and distance over 50 realizations for GMSQ: random walks on the generalized MSQ networks and LEVY: Lévy flighst on the lattice. Inset shows the case of spatially sparser distributions of communication requests for LEVY. The lines of GMSQ are duplicated.

gives more load for a ferry to store and carry the information of a large number of links. Table I shows the relation of the network size N and the buffer size for the same BR:0.2. In the generalized MSQ networks, the slopes of $P(l_{ij})$ in log-log plot correspond to $\mu = 1.414, 2.248, 2.692$, and 3.034 for $N = 500, 1000, 2000$, and 3000, respectively. These values of μ slightly differ from the example shown in Section IV because of using other area in the census data, however the obtained results are consistent.

Figure 14(a)(b) shows the average time step and distance until a ferry of free mode encounters a node s. The time step increases as the value of μ is larger, since the length of movement in one hop tends to be small on the dense network and on a Brownian motion. By the above effect, the distance also increases as the value of μ is larger. In the Lévy flights, the part of extremely large distance for $\mu < 2$ is due to a ballistic motion, and the shortest distance is obtained around $\mu = 2.4$. Remember that such a U-shape graph of μ vs. distance is obtained as the inverse U-shape graph of μ vs. scaled efficiency in Figure 8. Here, for the Lévy flights of $\mu = 1.6 \sim 2.8$, the positions of s-d nodes are set at the nodes of the generalized MSQ networks of a large size $N = 3000$, while they are set at the nodes of them of a small size $N = 500$ in Inset. We call these positions POS-N3000 and POS-N500, corresponding to spatially dense and sparse distributions of communication requests. Figure 15(a)(b) shows the average time step and distance until a ferry finds a path between s and d from the encounter with $REQ(s,d)$. Similar behavior to Figure 14(a)(b) is obtained, although there are dependences on the BR; The time step and distance become shorter, as the BR is larger. In both



(a) Time

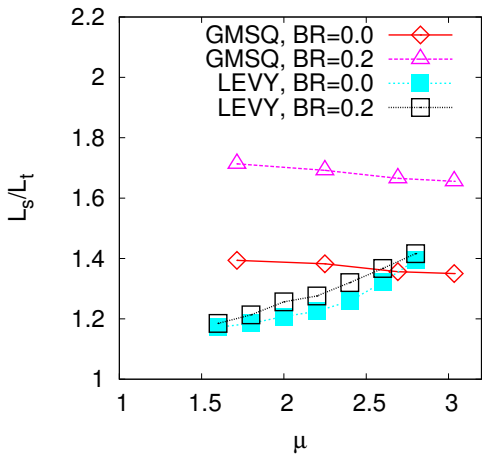


(b) Distance

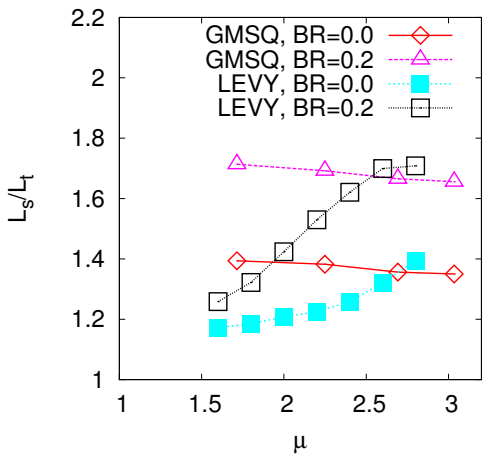
Figure 15. Average search time and distance over 50 realizations for GMSQ: random walks on the generalized MSQ networks and LEVY: Lévy flights on the lattice. Inset shows the case of spatially sparser distributions of communication requests for LEVY. The lines of GMSQ are duplicated.

Figures 14 and 15, we emphasize that the cases of GMSQ show shorter time and distance than the cases of LEVY. The most efficient size is $N = 500$ corresponding to the smallest μ plotted at the left end in the figures. We note that, depending on the situation of movements of ferries and locations of destination nodes, the links memorized in the buffer of a ferry maybe not work well, since hundreds of times are spent for the encounter and the search, especially for a large μ.

The ratio of the path lengths obtained by the routing and by the shortest distance on the generalized MSQ network is between 1.1 and 1.8 as shown in Figure 16(a)(b). For the reason that the case for BR:0.2 is worse than the case for BR:0 without cooperations between ferries, the higher ratio of L_s/L_t is caused from less information used for finding a path as shown in Figure 17(a)(b). In addition, Figure 17(a)(b) shows that the average number of used links in the subgraph for calculating a path is between 20% and 40% of the total number of links. There is a trade-off: In Figure 17(a) for BR:0.2, GMSQ is slightly better than LEVY for using less information, however the path length is longer as shown in Figure 16(a). Figure 18(a)(b) shows that the average number of requests carried by a ferry is less than 6 in the cases of GMSQ, and smaller than that in the cases of LEVY. Thus, in the cases of GMSQ for BR:0.2, a routing path is obtained at most 1.4 times longer than the shortest distance by using only partial information about 20% of the total number of links, in average.



(a) POS-N500

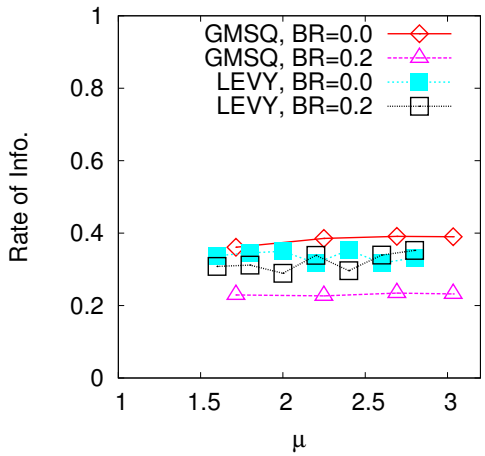


(b) POS-N3000

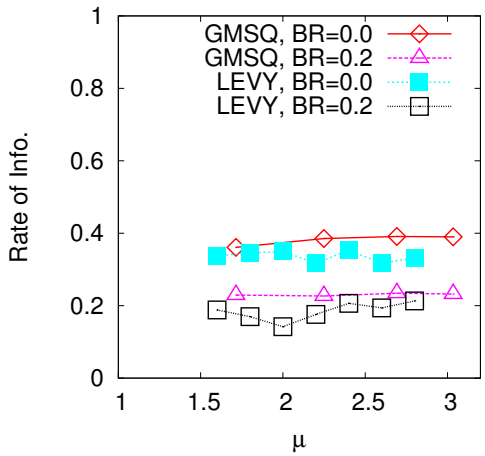
Figure 16. Ratio of the path length L_s obtained by the routing and the shortest distance path L_t. For LEVY, the spatially (a) sparse and (b) dense distributions of communication requests are generated on the nodes of the generalized MSQ networks for $N = 500$ and $N = 3000$, respectively. Here, L_s is between 80 and 120, therefore L_t is in the same order.

VI. Conclusion

We have considered a scalable self-organized geographical network by iterative divisions of rectangles for load

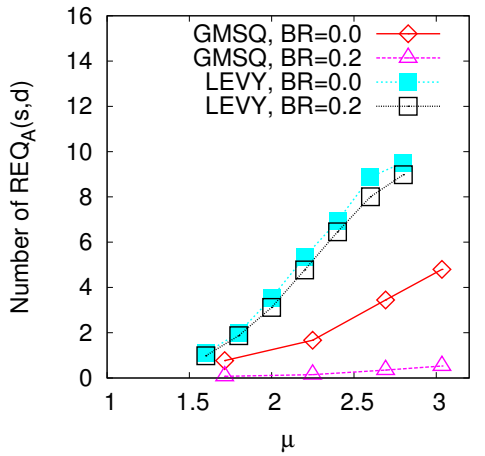


(a) POS-N500

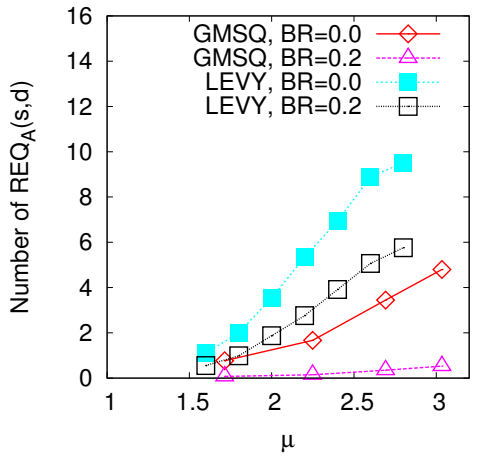


(b) POS-N3000

Figure 17. Rate of partial information: (the average number of used links in the subgraph for calculating a path) divided by (the total number of links). For LEVY, the communication requests are generated on the spatially (a) sparse and (b) dense networks.



(a) POS-N500



(b) POS-N3000

Figure 18. Average number of communication requests carried by a ferry. For LEVY, the requests are generated on the spatially (a) sparse and (b) dense networks.

balancing of nodes in the adaptive change of their territories according to the increase of communication requests. In particular, the spatially inhomogeneous distributions of population and the corresponding positions of nodes are important. For the proposed networks [1], [10], we have investigated the search efficiency in the destructive case [19] with new creations of target after the detections, and shown that the α-random walks [17], [41] on the networks within a small size have higher search efficiency than the Lévy flights known as the optimal strategy [18], [19] for homogeneously distributed targets on the square lattice with periodic boundary conditions. One reason for the better performance is the anisotropic covering of high population areas.

Furthermore, we apply the good property to the decentralized routing by cooperative message ferries. The key point is also the adaptation of network structure to the spatial distributions of source and destination which are inhomogeneous according to a population data. As the merit, by using only a simple protocol based on random walks, the naturally embedded fractal-like sparse structure contributes to the search of targets and to the find of a path efficiently in such a realistic situation of spatially distributed communication request.

However, we must take care of the size. Our method on the generalized MSQ networks within a small size shows better performance in both time and distance than the Lévy flight version using only partial information of links. For the scale up issues, since the performance goes down as the size is larger, we should make various ideas to keep the appropriate size ($N \approx 500$), e.g., by the enhancement of processing power at a node, instead of distribution of load in a large size. The performance for both encounter and search

times will be improved by considering further methods of how to cooperate with ferries and nodes.

On the other hand, the message ferrying scheme is generally applicable to a temporal network, in which the positions of nodes and connections between them are changed in a short time. It is interesting to study such cases for the generalized MSQ networks with temporal disconnections. Since a human mobility pattern resembles to the Lévy flights, a good performance of the proposed cooperative routing will be expected for a temporal network that consists of multi-hop mobile communication equipments, although how to treat the temporal disconnections caused from node mobility will be one of the important issues. Instead of the message ferry scheme, it is worth to investigate the performance of other DTN routing methods on the generalized MSQ network. For more rigorous discussions about the performance, statistical tests [50] may be useful to clarify the applicability of the proposed method. The limitation for the applicability will also depend on future technologies of wireless devices.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for their valuable comments. This research is supported in part by Grant-in-Aide for Scientific Research in Japan, No.21500072 & No.25330100.

REFERENCES

[1] Y. Hayashi, "Adaptive Fractal-like Network Structure for Efficient Search of Targets at Unknown Positions," *Proc. of the 4th International Conference on Adaptive and Self-adaptive Systems and Applications (ADAPTIVE)*, pp.63-68, ISBN:978-1-61208-219-6, 2012.

[2] S. H. Yook, H. Jeong, and A.-L. Barabási, "Modeling the internet's large-scale topology," *PNAS*, Vol.99, No.21, pp.13382–13386, 2002.

[3] M. T. Gastner and M. E. J. Newman, "The spatial structure of networks," *Eur. Phys. J. B*, Vol.49, No.2, pp.247–252, 2006.

[4] R. Guimerà, S. Mossa, A. Turtschi, and L. Amaral, "The worldwide air transportation network: Anormalous centrality, community structure, and cities' global roles," *PNAS*, Vol.102, No.22, pp.7794–7799, 2005.

[5] R. Lambiotte, V. Blondel, C. de Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Dooren, "Geographical dispersal of mobile communication networks," *Physica A*, Vol.387, pp.5317–5325, 2008.

[6] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and its Applications (2nd Eds.)*, John Wiley & Sons, 1995.

[7] A.-L. Barabási, *Linked: The New Science of Networks*. Perseus, Cambridge, MA, 2002.

[8] M. Buchanan, *Nexus: Small Worlds and the Groundbreaking Theory of Networks*. W.W.Norton, New York, 2002.

[9] M.E.J. Newman, A.-L. Barabási, and D.J. Watts (Eds.), *The structure and dynamics of NETWORKS*. Princeton University Press, Princetion and Oxford, 2006.

[10] Y. Hayashi, "Rethinking of Communication Requests, Routing, and Navigation Hierarchy on Complex Networks -for a Biologically Inspired Efficient Search on a Geographical Space-," In T. Bilogrevic, A. Rezazadeh, and L. Momeni (Eds.), Networks -Emerging Topics in Computer Science, Chapter 4, pp. 67-88, iConcept Press, 2012.

[11] G. M. Viswanathan, M. G. E., Luz, E. P. Raposo, and H. E. Stanley, *The Physics of Foraging -An Introduction to Random Searches and Biological Encounters*, Cambridge University Press, 2011.

[12] I. Stojmeńović (Eds.), Handbook of Wireless Networks and Mobile Computing, John Wiley & Sons, 2002.

[13] A. Boukerche (Eds.), Handbook of Algorithms for Wireless Networking and Mobile Computing, Chapman & Hall, 2006.

[14] J. Urrutia, "Routing with Guaranteed Delivery in Geometric and Wireless Networks," *in Handbook of Wireless Networks and Mobile Computing, I. Stojmenović (Ed.)*, Chapter 18, John Wiley & Sons, 2002.

[15] W.-X. Wang, B.-H. Wang, C.-Y. Yin, Y.-B. Xie, and T. Zhou, Traffic dynamics based on local routing protocol on a scale-free network. *Phys. Rev. E*, Vol.73, pp.026111, 2006.

[16] B. Danila, Y. Yu, S. Earl, J. A. Marsh, Z. Toroczkai, and K. E. Bassler, "Congestion-gradient driven transport on complex networks," *Phys. Rev. E*, Vol.74, pp.046114, 2006.

[17] W.-X. Wang, C.-Y. Yin, C.-Y., G. Yan, and B.-H. Wang, "Integrating local static and dynamic information for routing traffic," *Phys. Rev. E*, Vol.74, pp.016101, 2006.

[18] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. E., da Luz, E. P. Raposo, and H. E. Stanley, "Optimizing the success of random searches," *Nature*, Vol.401, pp.911–914, 1999.

[19] M. C. Santos, G. M. Viswanathan, E. P. Raposo, and M. E. da Luz, "Optimization of random search on regular lattices," *Phys. Rev. E*, Vol.72, pp.046143, 2005.

[20] M. C. Santos, G. M. Viswanathan, E. P. Raposo, and M. E. da Luz, "Optimization of random searches on defective lattice networks," *Phys. Rev. E*, Vol.77, pp.041101, 2008.

[21] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, Vol.74, No.1, pp.47–97, 2002.

[22] N. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, Vol.45, No.2, pp.167–256, 2003.

[23] R. Xulvi-Brunet and I. Sokolov, "Evolving networks with disadvantaged long-range connections," *Phys. Rev. E*, Vol.66, pp.026118, 2002.

[24] S. S.Manna and P. Sen, "Modulated scale-free network in euclidean space," *Phys. Rev. E*, Vol.66, pp.066114, 2002.

[25] P. Sen and S. S. Manna, "Clustering properties of generalized critical euclidean network," *Phys. Rev. E*, Vol.68, pp.026104, 2003.

[26] A. K. Nandi and S. S. Manna, "A transition from river networks to scale-free networks," *New J. Phys.*, Vol.9, pp.30, 2007

[27] J. Wang and G. Provan, "Topological analysis of specific spatial complex networks," *Advances in Complex Systems*, Vol.12, No.1, pp.45–71, 2009.

[28] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, Vol.286, pp.509–512, 1999.

[29] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, Vol.393, pp.440–442, 1998.

[30] Z. Zhang, S. Zhou, Z. Su, T. Zou, and J. Guan, "Random sierpinski network with scale-free small-world and modular structure," *Euro. Phys. J. B*, Vol.65, pp.141–148, 2008.

[31] T. Zhou, G. Yan, and B.-H. Wang, "Maximal planar networks with large clustering coefficient and power-law degree distribution," *Phys. Rev. E*, Vol.71, pp.046141, 2005.

[32] Z. Zhang and L. Rong, "High-dimensional random apollonian networks," *Physica A*, Vol.364, pp.610–618, 2006.

[33] J. P. K. Doye and C. P. Massen, "Self-similar disk packings as model spatial scale-free networks," *Phys. Rev. E*, Vol.71, pp.016128, 2005.

[34] L. Wang, F. Du, H. P. Dai, and Y. X. Sun, "Random pseudofractal scale-free networks with small-world effect," *Eur. Phys. J. B*, Vol.53, pp.361–366, 2006.

[35] H. D. Rozenfeld, S. Havlin, and D. ben Avraham, "Fractal and transfractal scale-free nets," *New J. of Phys.*, Vol.9, pp.175, 2007.

[36] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Pseudofractal scale-free web," *Phys. Rev. E*, Vol.65, pp.066122, 2002.

[37] R. Albert and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, Vol.406, pp.378–382, 2000.

[38] Y. Hayashi, "Evolutionary construction of geographical networks with nearly optimal robustness and efficient routing properties," *Physica A*, Vol.388, pp.991–998, 2009.

[39] Y. Hayashi and Y. Ono, "Geographical networks stochastically constructed by a self-similar tiling according to population," *Phys. Rev. E*, Vol.82, pp.016108, 2010.

[40] M. I. Karavelas and L. J. Guibas, "Static and kinetic geometric spanners with applications," In *Proc. of the 12th ACM-SIAM Symposium on Discrete Algorithms*, pp. 168–176, 2001.

[41] Y. Hayashi and Y. Ono, "Traffic properties for stochastic routing on scale-free networks," *IEICE Trans. on Communication*, Vol.E94-B(5), pp.1311–1322, 2011.

[42] J. D. Noh and H. Rieger, "Random walks on complex networks," *Phys. Rev. Lett.*, Vol.92, No.11, pp.118701, 2004.

[43] H. Shah, "Routing Enhancement Specific to Mobile Environment Using DTN," *International Journal of Computer Theory and Engineering*, Vol.3, No.4, pp. 537-542, 2011.

[44] R. J., D'Souza and J. Jose, "Routing Approaches in Delay Tolerant Networks: A Survey," *International Journal of Computer Applications*, Vol.1, No.17, pp. 8-14, 2010.

[45] W. Zhao, and M. H. Ammar, "Message Ferrying: Proactive Routing in Highly-partitioned Wireless Ad Hoc Networks," *Proc. of the 9th IEEE Workshop on Future Trends of Distributed Computing Systems (FTDCS)*, pp.308-314, 2003.

[46] W. Zhao, M. H. Ammar, and E. Zegura, "A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks," *Proc. of the 5th ACM International Symposium on Mobile Ad Hoc Networking (MOBIHOC)*, pp.187–198, 2004.

[47] W. Zhao, M. H. Ammar, and E. Zegura, "Controlling the Mobility of Multiple Data Transport Ferries in a Delay-Torerant Network," *Proc. of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pp.1407–1418, 2005.

[48] M. Shin, S. Hong, and I. Rhee, "DTN Routing Strategies using Optimal Search Patterns," *Proc. of the 3rd ACM Workshop on Challenged Networks (CHANTS)*, pp.27–32, 2008.

[49] N. Alon, C. Avin, M. Koucky, G. Kozma, Z. Lotker, and M. R. Tuttle, "Many Random Walks Are Fster Than One," *Combinatorics, Probability and Computing*, Vol.20, No.4, pp.481–502, 2011.

[50] A. Clauset, C. R. Shalizi, M .E. J. Newman, "Power-Law Distributions in Empirical Data," *SIAM Review*, Vol.51, pp.661–703, 2009.