An Information Model for Geographic Greedy Forwarding in Wireless Ad-Hoc Sensor Networks

Zhen Jiang
Computer Science Department
West Chester University
West Chester, PA 19383, USA
zjiang@wcupa.edu

Junchao Ma, Wei Lou
Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
{csjma, csweilou}@comp.polyu.edu.hk

Jie Wu
Department of Computer Sci. & Eng.
Florida Atlantic University
Boca Raton, FL 33431, USA
jie@cse.fau.edu

***Abstract*—In wireless ad-hoc sensor networks, an important issue often faced in geographic greedy forwarding routing is the "local minimum phenomenon" which is caused by deployment holes and blocks the forwarding process. In this paper, we provide a new information model for the geographic greedy forwarding routing that only forwards the packet within the so-called request zone. Under this new information model, the hole and its affected area are identified easily and quickly in an unsafe area with a labeling process. The greedy forwarding will be blocked if and only if a node inside the unsafe area is used. Due to the shape of the request zone, an unsafe area can be estimated as a rectangular region in the local view of unsafe nodes. With such estimate information, the new routing method proposed in this paper will avoid blocking by holes and achieve better performance in routing time while the cost of information construction is greatly reduced compared with the best results known to date.**

Keywords: Distributed algorithm, information model, routing, wireless ad-hoc sensor networks.

I. Introduction

Geographic greedy forwarding [1], [2], as a simple, efficient and scalable strategy, is the most promising routing scheme in wireless ad-hoc sensor networks (WASNs). In such a scheme, the routing path from the source to the destination is determined by the forwarding node selection at each intermediate node in a fully-distributed manner. The packet is forwarded hop by hop along such a path. An important challenge often faced in geographic greedy forwarding in WASNs is the "local minimum phenomenon" [3] which is caused by deployment *holes* where the forwarding process is blocked at a node called *stuck node*. The occurrence of hole can be caused by many factors, such as sparse deployment, physical obstacles, node failures, communication jamming, power exhaustion, and animus interference [3]–[7].

To mitigate the local minimum issue, Greedy-Face-Greedy (GFG) [8], Greedy Perimeter Stateless Routing (GPSR) [9], and Greedy-Other-Adaptive-Face Routing (GOAFR) [10] are currently the most popular methods. When routing process gets stuck at an intermediate node, it will start a perimeter routing phase where the packet is routed by the "right-hand rule" counter-clockwise along a face of the planar graph that represents the same connectivity as the original network, until it reaches a node that is closer to the destination than that stuck node. After that, the routing returns to the greedy forwarding phase. In recent work [11]–[13], such a routing scheme is proved to guarantee delivery in any arbitrary planar graph. However, without enough shape information of the holes, such a routing may use a long detour path in the perimeter routing, compared with the shortest path to the destination [14].

In this paper, we propose a simple and efficient method to achieve a much shorter path than any traditional one in the geographic greedy forwarding routing. Inspired by an early work of *safety level* [15] and its extensions [16], we develop an information model to identify the affected area of each hole as an unsafe area. Many nodes near the stuck nodes are also affected because their successors are all stuck nodes. Thus, all such nodes, including both the stuck nodes and some of their neighboring nodes, are unsafe. By limiting the greedy forwarding within the request zone in [2], the shape of each unsafe area can be estimated as a rectangular region, which has a simple structure for easy construction and maintenance. Moreover, by considering the relative locations of the source and the destination, such an unsafe area is optimal; that is, the local minimum will occur if and only if an unsafe node is used in routing. In order to make the whole system scalable, such an information model is implemented in a fully distributed manner. With the estimated shape information collected and distributed only among a few unsafe neighbors, the corresponding routing can avoid the local minimum by always selecting safe node in forwarding phase. As a result, the routing performance will improve greatly while the cost of information model can be controlled to minimal.

The contributions of this paper are threefold.

- First, we provide a novel information model in which the forwarding hole and its affected nodes can be identified in an unsafe area and further be estimated as a rectangle. It is the first attempt to address the mutual impact of holes in blocking the routing. It is proved that such regions are optimal and their construction converges very quickly.
- Second, our new information model is applied to the routing. This is the first attempt to find the balance of tradeoff between the cost of information model and the routing adaptivity while pursuing better routing performance. The new information-based routing, denoted as SLGF, has the ability of precisely predicting the local minimum ahead, which distinguishes itself from the others.

- Third, we develop a simulator to show the substantial improvement of SLGF routing in term of the routing time and the cost of hole information model, compared with the best results known to date. The experimental results show that a cost-effective geographic greedy forwarding routing is created under our new information model.

The remainder of this paper is organized as follows: Section 2 introduces some necessary notations and preliminaries. Section 3 provides our new hole information model and its distributed construction process, including information collection, distribution, and storage. Some analytical results on the cost of this model are also provided. In Section 4, we introduce the SLGF routing. In Section 5, the experimental results are provided to show the performance improvement and the cost reduction in SLGF routing, compared with the best results known to date. Section 6 provides some background information and Section 7 concludes the paper and provides directions for future research.

II. Preliminary

With the assumption that all the sensors have the same communication range, a WASN can be represented by a simple undirected graph $G = (V, E)$, where V is a set of vertices including all the nodes and E is a set of undirected edges, each of which indicates two nodes are within the communication range of each other. $N(u)$ denotes the set of neighboring nodes of node u. Each node u has the location (x_u, y_u), simply denoted by $L(u)$. The location information can be discovered by having Global Positioning System (GPS) receivers [17] at some fixed nodes [18], or a mobile beacon node [19], or just relying on the relative coordinate system [20]. $\mid L(u) - L(v) \mid$ is the distance between two nodes u and v. $s(x_s, y_s)$ and $d(x_d, y_d)$ are the source and the destination nodes. $[x_1 : x_2, y_1 : y_2]$ represents a rectangle with four corners (x_1, y_1), (x_1, y_2), (x_2, y_2), and (x_2, y_1). In this paper, all the routing schemes are presented via their forwarding node selection at an intermediate node $u(x_u, y_u)$ along the path. Rectangle $[x_u : x_d, y_u : y_d]$ has both u and d at the opposite corners. It is also called the request zone of node u and denoted as $Z(u)$. When d is located in quadrant I (or the Northeast) of u, the routing is called type-I routing and the corresponding $Z(u)$ in quadrant I of u is called a type-I request zone (see in Figure 1 (a)). Similarly, we have type-II, III, and IV routings and their request zones (where $Z(u)$ locates in quadrants II, III, and IV of u, respectively), after a 90, 180, or 270 degree counter-clockwise rotation of the graph G. Note that the forwarding for one packet may experience different types of request zones, when the relative position of d to u changes and d locates in different types of request zones (see in Figure 1 (b)). To increase the readability of this paper, we only focus on the discussion on type-I (i.e., Northeast type). To simplify the discussion, we describe all the schemes in a synchronous, round-based system. All the schemes presented in this paper can be extended easily to an asynchronous round based system. However, to make our schemes clear, we do not pursue the relaxation.

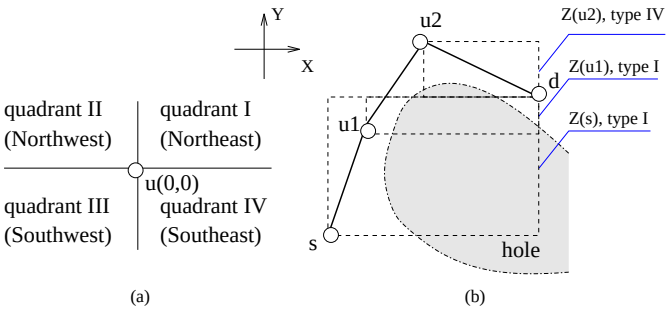


Fig. 1. Definition of different types of request zones.

In [21], the stuck node is identified where the packet can get the local minimum in greedy forwarding routing [8], [9], and then, a process called BOUNDHOLE is initiated to form a closed circle (also called a *boundary*). The region enclosed by the boundary is identified as the hole area. For each node u along the boundary, its successor node along the boundary in clockwise order is marked as the *downstream node* of u.

When the local minimum occurs, the routing must be at a stuck node u which is also a boundary node. The boundary-information-based routing in [21] will then route the packet downstream until it reaches a node v whose distance to d is closer than that of u. After that, the greedy forwarding phase can continue with the routing decision described in the LAR scheme 2 in [2]. The complete routing algorithm (also called GF routing) is shown in Algorithm 1.

Algorithm 1 Geographic greedy forwarding (GF) routing based on boundary information [21]

At the current node u (including the source s):

1: If $d \in N(u)$, forward the packet to d, and then stop.
2: Use the greedy forwarding to forward the packet to $v \in N(u)$ where $\mid L(v) - L(d) \mid < \mid L(u) - L(d) \mid$ [2].
3: If such a v does not exist, u is on the boundary. Repeat the perimeter routing to send the packet along the downstream until a node v, which satisfies the condition $\mid L(v) - L(d) \mid < \mid L(u) - L(d) \mid$, is reached.

As one of many traditional geographic greedy routings using "right-hand rule" policy [8]–[13] in the perimeter routing phase, the limited geographic greedy routing, denoted by LGF, selects the forwarding successor candidates within the request zone which is described in LAR scheme 1 in [2] as a rectangle in the corresponding quadrant. The successor node selection at the current node u in its perimeter routing phase is implemented by simply rotating the ray ud counter-clockwise until the first untried node $v \in N(u)$ is hit by the ray. The details of the LGF are shown in Algorithm 2.

Figure 2 (a) shows an example of LGF routing. First, the type-I greedy forwarding starts and the packet will be forwarded to node u_5. At node u_5, the routing cannot find the successor in the request zone $Z(u_5)$ so that the perimeter routing phase is conducted and the packet is forwarded to u_2.

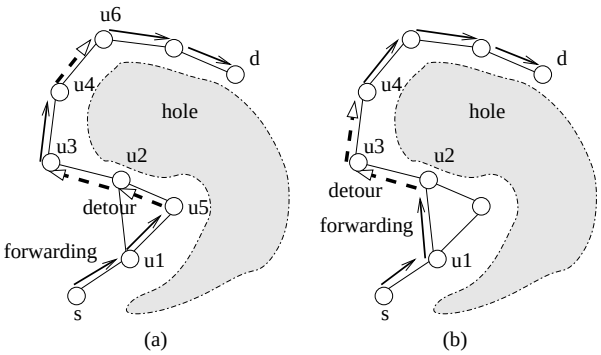


Fig. 2. (a) LGF routing (i.e., forwarding with LAR scheme 1). (b) GF routing [21] (i.e., forwarding with LAR scheme 2).

Algorithm 2 Limited geographic greedy forwarding (LGF) routing

At the current node u (including the source s):

1: If $d \in N(u)$, forward the packet to d, and then stop.
2: Determine the request zone $Z(u) = [x_u : x_d, y_u : y_d]$ and use the greedy forwarding to forward the packet to $v \in (N(u) \cap Z(u))$ [2].
3: If such a v does not exist, send the packet in the perimeter routing by the "right-hand rule" policy [8]–[13].

When the routing reaches node u_6, the request zone is of type IV. Starting from u_6, the packet will be forwarded in type-IV greedy forwarding until it reaches the destination d.

Compared with GF routing in [21], the forwarding in LGF routing is more straightforward. Without the effect of holes, LGF routing will achieve a shorter path. However, it has fewer number of different paths (see the only selection of u_5 at u_1 in Figure 2 (a)). Thus, if there is a hole ahead, LGF routing has fewer opportunities to continue the greedy forwarding phase and may experience more perimeter routing phases (compare the routing in Figure 2 (b)). As a result, it may have a longer routing path than GF routing in the presence of holes.

In the next section, we will present our hole information model for LGF routing. The information-based LGF routing can achieve better performance than GF routing in terms of the length of routing path (i.e., the speed of routing). In this way, we show the impact value of our new information model.

III. Safety Model for LGF Routing

In this section, we introduce the concept of safe/unsafe status. The hole and its nearby connected unsafe nodes will form an unsafe area. Then, we prove that such an unsafe area is optimal for LGF routing. In the local view of each unsafe node, the unsafe area is estimated as a rectangular region. We provide a distributed construction process to collect and distribute the safety information and the estimated shape information in the entire network.

In LGF routing, the perimeter routing phase starts when the current node u has no successor candidate inside its request zone; that is, the local minimum occurs. In our new information model, the nodes are labeled as unsafe nodes if using them and only using them will cause a local minimum. Due to the types of request zones, there are four different types of safe/unsafe statuses. Type-I safe/unsafe status is identified for type-I routing using type-I request zone. Respectively, we have safe/unsafe statuses of type-II, III, and IV. The definition of safe/unsafe status of each type and the corresponding labeling processes are shown as follows. After the labeling process, each node will be labeled with a 4-tuple (NE, NW, SW, SE) where NE stands for the safe/unsafe status of type-I. NW, SW, and SE are the ones of type-II, III, and IV, respectively.

Definition 1 (labeling process): *Initially, each healthy node $u(x_u, y_u)$ sets its status 4-tuple (NE, NW, SW, SE) to (s, s, s, s), where symbol "s" (or "u") stands for the safe (or unsafe) status and NE, NW, SW, and SE are the safe/unsafe statuses of type-I, II, III, and IV, respectively. Any status, say NE, will be changed to unsafe if there is no type-I safe neighbor in the type-I request zone; that is, $\{v(x_v, y_v) \mid v \in N(u) \wedge v$ has label $(s, -, -, -) \wedge x_u < x_v \wedge y_u < y_v\} = \emptyset$, where symbol "$-$" stands for either safe or unsafe status.*

Based on Definition 1, a node u may change its safe status of one type, say type-I, to unsafe status; that is, a hole in the type-I request zone is detected. This change may also affect its neighbors' safety information and contribute further changes. The hole and all nearby connected type-I unsafe nodes form a so-called *type-I unsafe area*. According to the type of unsafe nodes contained, other types of unsafe areas, type-II, III, and IV unsafe areas, form respectively. We assume that all of the communication actions occur inside a region that is called the *interest area*. In our labeling process, any node out of the interest area will always keep its status tuple as (s, s, s, s). Moreover, a node inside the interest area will always keep one certain safe status, for instance, NE, when its communication disk in the corresponding quadrant intersects the edge of interest area. Thus, the edge of interest area will not affect the label of nodes inside. The following theorem shows that the unsafe area is optimal for a LGF routing.

Theorem 1: *For any hole that can block an LGF routing with certain type, say type-I, using and only using a type-I unsafe node in the corresponding unsafe area will definitely cause the local minimum.*

Proof: For the last hole that can block a type-I LGF forwarding, s is in its Southwest and d is in its Northeast. Moreover, d cannot be an unsafe node in its unsafe area. Obviously, when such a local minimum occurs, the LGF routing has reached a type-I unsafe node along the perimeter of the hole. On the other hand, if a greedy forwarding path $u \rightarrow u_1 \rightarrow u_2 \rightarrow \cdots \rightarrow u_i = d$ exists, we can find an unsafe node u_j that its successor u_{j+1} is safe. According to the definition of unsafe status of u_j, u_{j+1} cannot be in its type-I request zone and cannot be the successor of u_j. ■

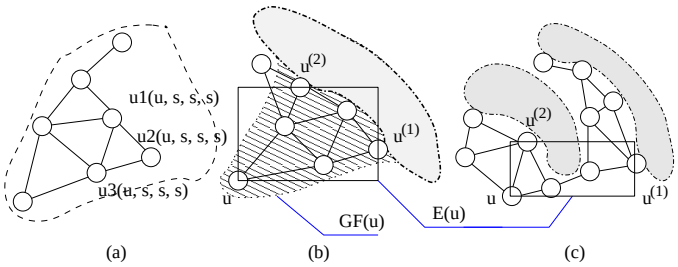


Fig. 3. Labeling process for type-I unsafe nodes. (a) stabilized safety information, (b) $GF(u)$ & $E(u)$, and (c) $E(u)$ with mutual impact of holes.

A sample of the labeling process is shown in Figure 3 (a). Initially, all nodes will set their status 4-tuple as (s, s, s, s). Only those nodes within the interest area will change their safety statuses. In the first round, nodes u_1 and u_2 will change its NE status to unsafe. In the second round, this unsafe status change will cause the change of the NE status at node u_3. The NE status of nodes u_1, u_2, and u_3 indicates a hole area in the direction of Northeast. It is noted that u_4 will keep its safe statuses because its communication disk in each quadrant intersects the edge of interest area.

Assume node u is a type-I unsafe node. The part of the type-I unsafe area in quadrant I describes the effect of the hole that blocks the forwarding of type-I routing from node u in the Northeast direction. We introduce the concept of *greedy forwarding region* of type-I unsafe node u, $GF_{NE}(u)$. Such a region includes all the unsafe nodes that can be reached from u by forwarding in the Northeast direction. Respectively, we have $GF_{NW}(u)$, $GF_{SW}(u)$, and $GF_{SE}(u)$ to indicate the greedy forwarding regions, which include the unsafe nodes in type-II, III, and IV unsafe areas reached by greedy forwarding in the corresponding direction. To simplify the discussion, we can only focus on $GF_{NE}(u)$, simply denoted by $GF(u)$.

For any node $v \in GF(u)$, we can always find a path $v_0(= u)$, v_1, v_2, $\cdots$, $v_n(= v)$, such that v_i $(0 \leq i \leq n-1)$ is unsafe and v_{i+1} is inside the type-I request zone of v_i. A sample of $GF(u)$ is shown in Figure 3 (b). Rotate a ray from u scanning $GF(u)$, counter-clockwisely. We denote that $u^{(1)}$ and $u^{(2)}$ are the farthest nodes that can be reached on the first and the last greedy forwarding paths. When the forwarding in type-I routing reaches node u, $u^{(1)}$ or $u^{(2)}$ can be used as the bound to detour around the hole. In this way, a short detour path is guaranteed and many unnecessary detours inside the unsafe area can be avoided. Therefore, from the view of node u, the shape of unsafe area H can be estimated as $H \cup [x_u : x_{u^{(1)}}, y_u : y_{u^{(2)}}]$. Furthermore, for a type-I forwarding routing, the shape of unsafe area can simply be represented by $E(u)$: $[x_u : x_{u^{(1)}}, y_u : y_{u^{(2)}}]$.

Individually, each unsafe node u will have its own estimated shape information of the related unsafe area. To collect and distribute such information, we have the following implementation. When u has no neighbor in the request zone, according to the definition, $u^{(1)} = u^{(2)} = u$. For the rest of cases, the location information of $u^{(1)}$ and $u^{(2)}$ is collected as well as the propagation of unsafe status. It is because that each node w along that greedy forwarding path from u to $u^{(1)}$ will have $w^{(1)} = u^{(1)}$. Node u can collect the location information of $u^{(1)}$ from its neighbor along that path, i.e., the first type-I unsafe neighbor hit by a ray from u when scanning the type-I requst region in counter-clockwise order. Similarly, we have a path from u to $u^{(2)}$ for the update of $u^{(2)}$ at u.

Figure 3 (b) shows the forwarding region of the type-I unsafe node u and the corresponding farthest reachable nodes $u^{(1)}$ and $u^{(2)}$. Then, the shape of the hole area in the Northeast is estimated as $E(u)$: $[x_u : x_{u^{(1)}}, y_u : y_{u^{(2)}}]$. Figure 3 (c) shows the sample of $E(u)$ in the case when u is mutually affected by intertwined holes. In the following theorem, we show that the convex rectangle $E(u)$ is equivalent to the unsafe area H for the routing at node u.

Theorem 2: *The greedy forwarding from node u in LGF routing, say type-I, will be blocked iff any node inside the estimated unsafe area of an unsafe node u, $E(u)$ $[x_u : x_{u^{(1)}}, y_u : y_{u^{(2)}}]$, is used.*

Proof: Any type-I forwarding will be blocked at a type-I unsafe node. Such a node u is the only node inside $E(u)$.

For any unsafe node u, if a node w inside $E(u)$ can be reached in the type-I forwarding from u, $w \in GF(u)$ and w is type-I unsafe. Using w will cause a local minimum according to Theorem 1. ■

Algorithm 3 shows the details of the construction process in our safety model. In such a process, the safety status and the estimated shape information are collected and distributed via information exchanges among neighbors. Such an exchange is implemented by broadcasting such information of a node that newly changes its safety status to all its neighbors.

Algorithm 3 Information construction

1: Each healthy node is initially labeled as a safe node.
2: For each safe node, change one of its status to unsafe, say NE, if there is no type-I safe neighbor within quadrant I.
3: For an unsafe node, say type-I unsafe node, set $u^{(1)} = u^{(2)} = u$ if $N(u) \cap Z(u) = \phi$. Otherwise, $u^{(1)} = v_1^{(1)}$ and $u^{(2)} = v_2^{(2)}$, where v_1 and v_2 are the first and the last type-I unsafe neighbors hit by a ray from u when scanning the type-I request region in counter-clockwise order.
4: Repeat steps 2 and 3 until no safe node changes its status.

Theorem 3: *For each unsafe area that only contains one hole, Algorithm 3 converges within $\frac{n}{2}$ rounds where n is the number of boundary nodes [21] encircling it.*

Proof: Assume that R is a rectangular region that exactly covers the final stabilized unsafe nodes. Thus, we can find every converging path whose length is shorter than half of the perimeter of R. A converging path is a sequence of nodes $u_1 \rightarrow u_2 \rightarrow \cdots \rightarrow u_k$ where u_i $(1 < i \leq k)$ obtains its unsafe status at round i. The number of rounds needed in a converging process is the maximum length of the converging

path. Denote L_c, L_R, and L_H as the length of the converging path, the length of the perimeter of region R, and the length of perimeter of the only hole in R, respectively. Due to the constitution of R in Algorithm 3, $L_c < \frac{L_R}{2}$ and $L_R \leq L_H$. That is, $L_c < \frac{L_H}{2}$. Because length of any boundary is no less than L_H, the nodes along the converging path are no more than half of the nodes along a boundary in probability; i.e., Algorithm 3 converges within $\frac{n}{2}$ rounds. ■

As indicated in [21], the lower bound and the upper bound of converging speed of boundary construction is n and $6 \times n$, respectively. The above theorem shows that the construction in our safety model converges much faster than the one in boundary model in [21] in the networks without intertwined holes; i.e., when each unsafe area contains only one hole. Their converging speeds in the networks with intertwined holes are compared in the experimental results in Section V.

We compare the boundary model [21] with our safety model in terms of the cost of construction process as the follows:

1) A hole can be encircled by several boundaries initiated by different stuck nodes, and many unnecessary nodes are only included in a boundary to complete the circle. It is proved in Theorem 1 that unsafe areas are optimal for the LGF routing by considering the mutual impact of holes in blocking the routing; that is, LGF routing will cause a local minimum iff a node included in the unsafe area is used. Due to different types of LGF routing, there are four types of unsafe nodes and unsafe areas. However, the number of unsafe nodes with any of its four safety statuses unsafe is still less than that of boundary nodes.
2) A boundary may be concave. The shape of the unsafe area has been optimized to a rectangle in the local view of each unsafe node. As a result, the construction process under our safety model is easier and faster than the one under the boundary model.
3) Each boundary node stores only the location of its downstream node. Compared with other models that require to specify the border of a polygon, the boundary model has the least storage requirement known to date. Each unsafe node only stores the location of the opposite corner of rectangle $E(u)$. Therefore, the storage requirement under the new information model is still kept minimal.
4) Building the boundary requires the identification of the stuck node and the downstream node. These identification processes are complex. Under the safety model, no calculation is needed. That is, the energy used for computation can be conserved greatly.

IV. Safety-Information-Based LGF Routing

In this section, we apply the estimated shape information and the safety status information to LGF routing. The new routing is called safety-information-based LGF routing, or simply, SLGF routing.

Basically, for a routing decision at the current node, the safe neighbor within the request zone is always preferred.

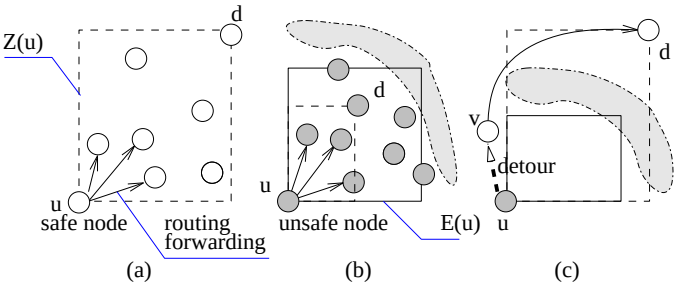


Fig. 4. Different routing cases in SLGF routing. (a) Forwarding with $Z(u)$, (b) forwarding when $d \in$ an unsafe area, and (c) perimeter routing by "right-hand rule".

Whenever an unsafe neighbor exists, the unsafe area in neighborhood is detected. With the estimated shape information stored in unsafe nodes, the routing will avoid entering an affected hole area and achieve a more straightforward routing path. When the destination is inside an unsafe area, the corresponding estimated shape information will guide the routing to route into that area. We also consider the routing when the source is already inside an unsafe area. In that case, the estimated shape information of the neighbors will help the routing retreat from that unsafe area, and then, find a path to the destination. In summary, the routing will use the estimated shape information stored at the unsafe nodes to conduct the routing phases in the following order: (1) forwarding to an unsafe node when d is inside the corresponding unsafe area (also called enforced forwarding), (2) forwarding to a safe node within the request zone (also called safe forwarding), (3) detouring to a safe node by "right-hand rule" (perimeter routing), and (4) retreating from an unsafe area.

We will now explain how to use these phases in a type-I routing. From a type-I safe node, we can always find a way to approach to the destination. Thus, any of its safe neighbors in $Z(u)$ can be a candidate of the successor node in the greedy forwarding process (safe forwarding phase). A sample is shown in Figure 4 (a). When the routing at u knows at least one type-I unsafe node in neighborhood, including u itself, say node w with the estimated shape information $E(w)$: $[x_w : x_{w^{(1)}}, y_w : y_{w^{(2)}}]$, the routing will first check if $d \in E(w)$. In that case, the routing has to approach to node d among all the unsafe nodes (enforced forwarding phase). The LGF routing that does not need safety information will be applied to find a closer node in the area $E(w)$ until the destination is reached. A sample case is shown in Figure 4 (b). When both above forwarding phases failed, the current node cannot be a safe node to d. In such a case, the detouring phase, and even retreating phase, will be conducted. First, rotate the ray of u in a counter-clockwise direction. After the last unsafe neighbor hit in $Z(u)$, the first node v safe to d hit by the ray (by the "right-hand rule") will be selected as the successor (perimeter routing phase). In other words, the successor node v will be searched in quadrants II, III, and IV in that respective order. Figure 4 (c) shows a sample of such a perimeter routing in finding the path around the holes. When the routing starts

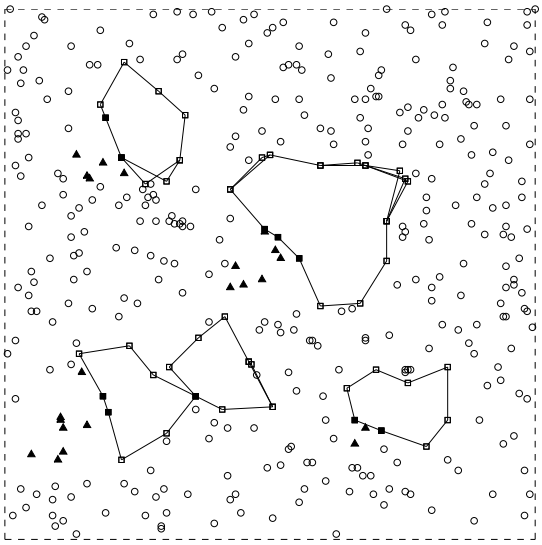
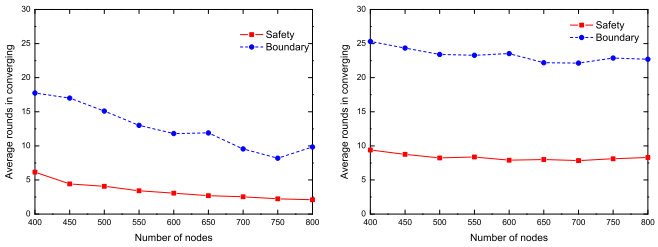
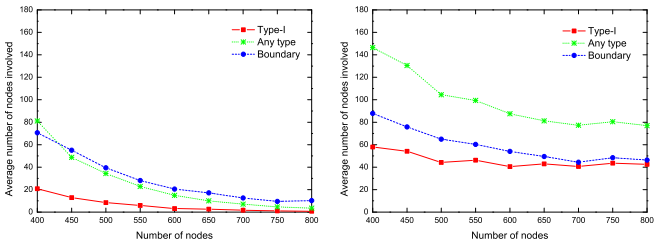


Fig. 5. The nodes involved in the information construction process under different information models. In the region, the white squares (□) are the boundary nodes, and the black triangles (▲) are the type-I unsafe nodes. Besides, the black squares (■) denote the nodes involved in both information models.



(a) Under IA model (b) under FA model

Fig. 6. Average number of rounds in converging for the boundary and the safety statuses.



(a) Under IA model (b) under FA model

Fig. 7. The number of nodes involved in boundary construction and safety information construction (in average).

from a source that is deep inside an unsafe area, the above perimeter routing phase may also fail because there is no safe neighbor around u. In such a case, the quickest way to leave such an unsafe area is to route in the opposite direction to d (retreating phase). For example, if the request zone is of type-I, the routing will go backwards to quadrant III. To ensure the routing can always get out of the unsafe area, a node that is not in the request zone will be selected as the successor node v. Such a retreating phase will continue until the above perimeter routing can be conducted successfully. The details of our information-based routing are shown in Algorithm 4.

Algorithm 4 Safety-information-based routing ($SLGF$) with the assumption that $x_u \leq x_d$ and $y_u < y_d$ (type-I)

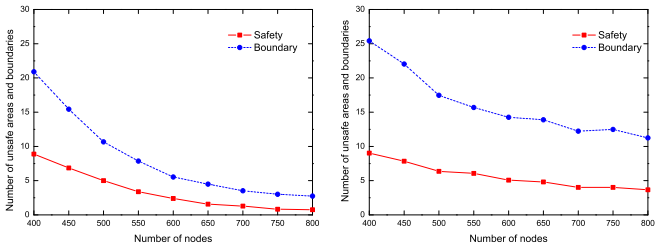
At the current node u (including the source s):

1: If $d \in N(u)$, forward the packet to d, and then stop.
2: Determine the routing type of each neighbor to d. Find an unsafe node $w \in (N(u) \cup \{u\})$ to d.
3: If $d \in E(w)$, apply the enforced forwarding, i.e, use the LGF routing to forward the packet to a node $v \in E(w)$.
4: If such a w does not exist, apply the safe forwarding, i.e., forward the packet to any safe node $v \in Z(u)$.
5: Otherwise, rotate the ray ud in counter-clockwise order to find the first node v safe to d, and then, send the packet to such a node v (i.e., the perimeter routing by the "right-hand rule").
6: If such a v is not found in step 5, apply the retreating, i.e., send the packet to any neighbor $v \in (N(u) \backslash Z(u))$ until the perimeter routing phase can be conducted.

V. Experimental Results

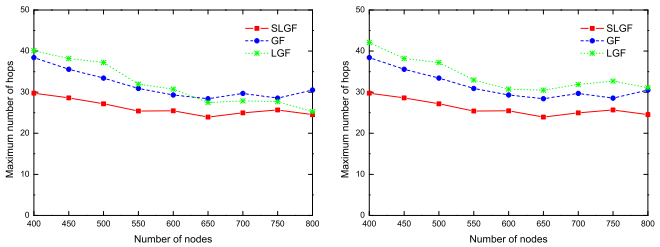
In this section, we study the average-case performance of the proposed information model and routing algorithms, using a simulator built in C++. The performance metrics used in the evaluation are the rounds in converging and the nodes involved in the hole identification process (i.e., cost of information model), and the hops and length of routing path.

In the simulations, nodes with a transmission radius of 20 meters are deployed to cover an interest area of 200m $\times$ 200m, under different deployment models. First, the nodes will be deployed uniformly. This is ideal model (denoted by IA), in which the hole is only caused by a sparse deployment. Usually, the size of a hole is very small. Secondly, we randomly set some forbidden areas inside interest area, where no nodes can be deployed. The forbidden areas, which may be irregular, are constructed to study the impact of larger holes on the proposed algorithms. Such a model is denoted by FA. We assume that the destination and the source are randomly selected in the interest area, including both safe sources and unsafe sources. Before we test the routing performance in routing time, within the interest area, boundary information [21] is constructed for GF routings, and safety information and estimated shape information are constructed for our SLGF routing. Figure 5 shows a sample of the nodes involved in different information construction processes. Then, we test the networks when the number of nodes in the interest area is varied from 400 to 800 in increments of 50. For each case, 100 networks are randomly generated, and the average routing performance over all of these randomly sampled networks is reported.



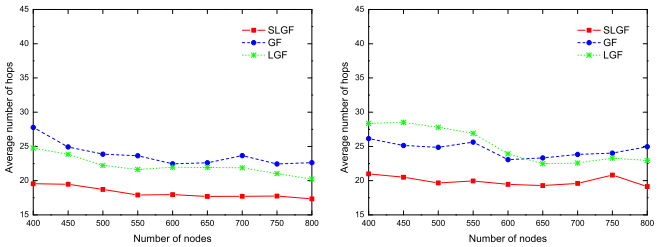
(a) Under IA model (b) under FA model

Fig. 8. Average number of type-I unsafe areas and average number of boundaries.

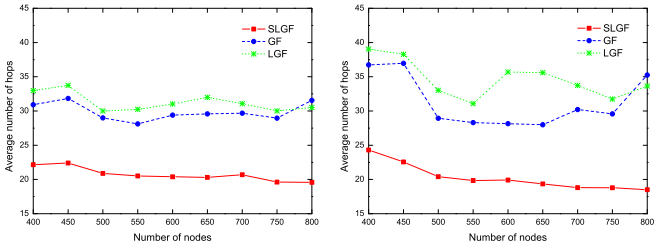


(a) Under IA model (b) under FA model

Fig. 9. Maximum number of hops of a GF, LGF, SLGF routing.



(a) Under IA model (safe source) (b) Under IA model (unsafe source)



(c) under FA model (safe source) (d) under FA model (unsafe source)

Fig. 10. Average number of hops of a GF, LGF, SLGF routing.

A. *Cost of information model*

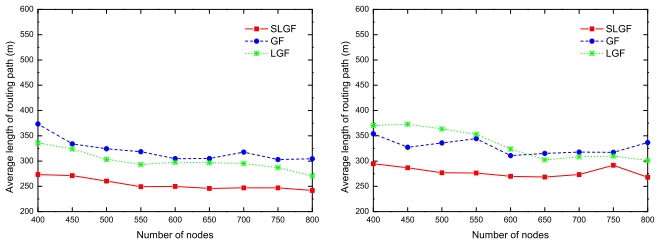
Figure 6 shows the average number of rounds in converging in different information constructions under both IA and FA models. Safety information construction converges faster and its implementation can be much easier. Thus, SLGF is much more practical for supporting fast responses to routing requests in WASNs. Because FA model has more large size holes, the converging speed of each construction will be slowed down.

Figure 7 shows the average number of type-I unsafe nodes involved in the information construction under both IA and FA models. As we mentioned earlier in Section III, only a node with newly updated safety status needs to broadcast the information. Thus, such a number implies the number of 1-hop broadcasting needed for type-I information construction. That is, the results in Figure 7 also show the message complexity of Algorithm 3. Type-II, III, and IV unsafe nodes have similar results. A node having any of its four safety statuses unsafe is called an "any type" unsafe node. The average number of any-type-unsafe nodes is also shown in Figure 7. The results show that fewer nodes are involved in the safety information construction than that in boundary construction (denoted by "boundary" in Figure 7), under IA model. As the node density increases, the number of nodes involved will decrease due to the smaller sizes of the holes. Obviously, when the size of holes becomes larger, under either the IA model or the FA model, more holes mutually impact each other and play the blocking role together; That is, more nodes, especially under the FA model, will be involved in safety information construction while such a mutual relation cannot be described in boundary nodes.

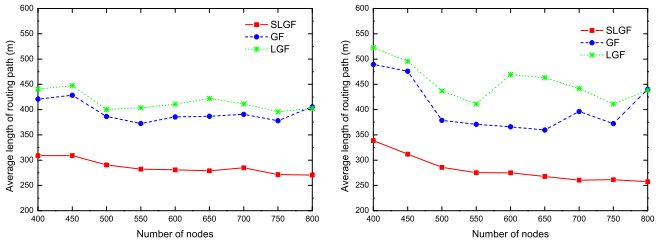
Figure 8 shows that the number of boundaries is always larger than the number of unsafe areas identified. Therefore, when new holes occur in the networks under either deployment model, the information maintenance for boundaries is much more difficult than that under our safety model.

B. *Hops and length of routing*

Figure 9 shows the upper bound of the number of hops of routing path. Respectively, Figure 10 shows the average number of hops of routing path. As we mentioned earlier in Section II, when holes exist in the networks, LGF routing may experience more perimeter routing phases than GF routing, because its forwarding adaptivity is limited and it will experience more blocking cases that have no forwarding node to use. As a result, LGF needs more hops. With the information used under our safety model, the routing can predict the holes ahead and avoid being blocked. In this way, SLGF routing can keep the forwarding direction in many cases and require the fewest number of hops in detour. It is because both GF and LGF routings have no information of any hole ahead and they can only route blindly. As shown in Figure 10, SLGF routing reduces the number of hops of routing significantly by approximately 25 percent as compared to GF, and about 35 percent as compared to LGF. When more large size holes occur under FA model, the above property still holds. In WASNs, the packet is forwarded hop by hop along the path. Reducing the number of hops can reduce end-to-end delay and furthermore support quick responses to routing requests. Figure 11 shows the corresponding length of entire routing path on average. These results prove the routing under our new information model can always achieve shorter path and conserve more energy used in data transmission. In Figure 10 and Figure 11, the comparison of the routings from unsafe sources with the ones from safe sources is provided. The results show that a routing with unsafe source needs few hops to retreat from hole area and can achieve nearly the same performance as it starts

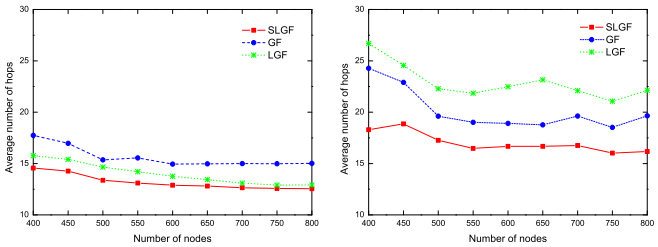


(a) Under IA model (safe source) (b) Under IA model (unsafe source)

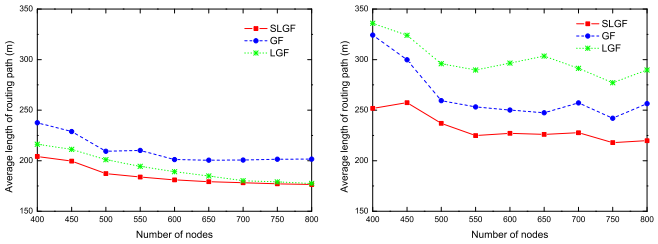


(c) under FA model (safe source) (d) under FA model (unsafe source)

Fig. 11. Average length of a GF, LGF, SLGF routing.



(a) Average number of hops (IA) (b) Average number of hops (FA)



(c) Average length (IA) (d) Average length (FA)

Fig. 12. Hops and length of the routing blocked by one single hole.

from a safe source. This proves the effectiveness of our safety model as it does not disable any possible communication and, indeed, improve the performance of them all. Figure 12 shows the corresponding results of routings without mutual impact of holes. Compared with the results in Figure 10 and Figure 11, it illustrates the consistent improvement of our safety model.

We summarize observations from the experimental results as follows. **(1)** The cost is reduced greatly under our safety model, compared with that of the boundary model in [21], in terms of the construction converging speed, the number of nodes involved, and the complexity of maintenance. By deploying more nodes or controlling the size of holes, the cost of both models will be reduced. By taking advantage of the broadcasting feature of wireless link, the message complexity of our safety model can be controlled to a certain level. **(2)** The proposed routing under safety model can always use fewer hops and achieve shorter path. As a result, the proposed routing has quicker response to routing request and conserve more energy used in data transmission.

VI. Related Work

In [22], some stuck nodes are identified as "dead ends". By removing the interference of holes, the detour in the perimeter routing phase will be more efficient. In [23], a local protocol produces short-cuts for the perimeter routing to bypass the holes. However, both routings cannot avoid the occurrence of every local minimum. In the approach presented in [24], the forwarding is guaranteed in the hyperbolic plane. However, whether every finite graph can be embedded in hyperbolic space is still up to questioning.

Recent work has focused on the use of hole area that contains the stuck nodes. In [21], the hole is detected at a node where packet can get local minimum in greedy forwarding routing [8], [9]. A process called BOUNDHOLE is initiated to form a closed circle (also called the boundary). The region enclosed by the boundary will be identified as the hole area. However, the identified hole area may not be convex. Without enough shape information of the hole area, the routing may not find a straight way to detour around the hole.

In [25], a stuck node contained in the convex hole area can be identified as a "float" node when the angle between its two adjacent neighbors exceeds a certain value ($210° - 230°$, suggested by the author). In [4], such an angle is called a turning angle. A corner that is located beside the boundary of a convex hole area is identified according to the values of its turning angles. Furthermore, the convex hole area is identified as the forbidden region via the corners. However, using the threshold of the turning angle is not precise enough to avoid every local minimum. In both approaches, even though many nodes may successfully forward the packet to the destination, they will be identified as stuck nodes and disabled from the consideration of routing decision as well as their forwarding paths. In many cases, the perimeter routing will be enforced while the forwarding path still exists. Therefore, it is important to minimize the shape of convex hole area so as to improve the routing path.

Another issue ignored in the existing approaches on determining the convex area is that the shape of a convex area is relative [26]; that is, its shape is different when the source and the destination of routing change their relative locations. In the existing schemes, the construction of convex hole areas always needs to re-do for each different routing case and costs a lot of time and energy.

VII. Conclusion

This paper is the first attempt to find the balance of the tradeoff between routing adaptivity and information model cost while pursuing better routing performance in WASNs. The paper is summarized as follows.

1) We optimize the shape of holes for LGF routing so that the corresponding information model proposed in this paper has an easy and quick construction, while its storage requirement does not increase. Under such an information model, the identified hole areas are optimal; that is, the information-based LGF routing does not need the perimeter routing detour whenever a forwarding path exists.
2) The new information model proposed in this paper has two uses: (a) the safe/unsafe status of a node is used to prevent a routing from entering into the affected area of a hole; (b) the estimated shape information E is used to predict the hole and achieve a straightforward path.
3) The information-based routing proposed in this paper mainly has three new parts: (a) route around the hole intelligently in a safe forwarding phase; (b) route away from the hole ahead (retreating phase); (c) route inside the unsafe area if needed (enforced forwarding phase). As a result, a more straightforward and shorter routing path can be achieved.
4) Such an improvement of routing in the presence of holes is the key to building a robust system in sparsely deployed WASNs, which has gained more attention recently. Moreover, the quickness of information construction makes our results more practical in networks where topology changes frequently, such as in mobile WASNs.

In our future work, we will extend our approach and search for a new balance point to increase the routing adaptivity so that fewer perimeter routing phases are needed and the routing path will be more straightforward and shorter. Also, we will conduct a further study on more accurate information for unsafe areas so that shorter paths can be achieved.

ACKNOWLEDGMENT

The work was supported in part by NSF grants ANI 0073736, CCR 0329741, CNS 0422762, CNS 0434533, EIA 0130806, CNS 0531410, and CNS 0626240, HKPU grants A-PG53, A-PH12, and Z09M, and grant PolyU-5236/06E.

REFERENCES

[1] S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, "A distance routing effect algorithm for mobility (DREAM)," *Proc. of the* 4^{th} *Annual ACM/IEEE International Conference on Mobile Computing and Networking (ACM/IEEE MOBICOM'98)*, 1998, pp. 76-84.

[2] Y. Ko and N. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," *Proc. of the* 4^{th} *Annual ACM/IEEE International Conference on Mobile Computing and Networking (ACM/IEEE MOBICOM'98)*, Oct. 1998, pp. 66-75.

[3] N. Ahmed, S. Kanhere, and S. Jha, "The holes problem in wireless sensor networks: A survey," *ACM Sigmobile Mobile Computing and Communication Review*, Vol. 9, No. 2, 2005, pp. 4-18.

[4] C. Chang, K. Shih, S. Lee, and S. Chang, "RGP: Active route guiding protocol for wireless sensor networks with obstacles," *Proc. of the* 3^{rd} *IEEE International Conference on Mobile Ad Hoc and Sensor Systems (IEEE MASS'06)*, Oct., 2006, pp. 367-376.

[5] K. Liu, N. Abu-Ghazaleh, and K. Kang, "Location verification and trust management for resilient geographic routing," *Journal of Parallel and Distributed Computing*, Vol. 67, No. 2, 2007, pp. 215-228.

[6] S. Olariu and I. Stojmenovic, "Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting," *Proc. of the* 25^{th} *Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM'06)*, April, 2006, pp. 1-12.

[7] X. Wu, G. Chen, and S. Das, "On the energy hole problem of nonuniform node distribution in wireless sensor networks," *Proc. of the* 3^{rd} *IEEE International Conference on Mobile Ad Hoc and Sensor Systems (IEEE MASS'06)*, Oct., 2006, pp. 180-187.

[8] P. Bose, P. Morin, and I. Stojmenovic, "Routing with guaranteed delivery in ad hoc wireless networks," *Proc. of the* 3^{rd} *International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, 1999, pp. 48-55.

[9] B. Karp and H. Kung, "GPSR: Greedy perimeter stateless routing for wireless sensor networks," *Proc. of the* 6^{th} *Annual ACM/IEEE International Conference on Mobile Computing and Networking (ACM/IEEE MOBICOM'00)*, August 2000, pp. 243-254.

[10] F. Kuhn, R. Wattenhofer, and A. Zollinger, "Worst-case optimal and average-case efficient geometric ad-hoc routing," *Proc. of the* 4^{th} *ACM International Sysmposium on Mobile Ad Hoc Networking and Computing (ACM MobiHoc'03)*, June 2003.

[11] H. Frey and I. Stojmenovic, "On delivery guarantees of face and comined greedy-face routing in ad hoc and sensor networks," *Proc. of the* 12^{th} *Annual ACM/IEEE International Conference on Mobile Computing and Networking (ACM/IEEE MOBICOM'06)*, 2006, pp. 390-401.

[12] Y. Kim, R. Govindan, B. Karp, and S. Shenker, "Geographic routing made practical," *Proc. of USENIX Symposium on Network Systems Design and Implementation*, May 2005.

[13] M. Nesterenko and A. Vora, "Void traversal for guaranteed delivery in geometric routing," *Proc. of the* 2^{nd} *IEEE International Conference on Mobile Ad Hoc and Sensor Systems (IEEE MASS'05)*, November 2005.

[14] M. Mauve, J. Widmer, and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *IEEE Network Magazine*, Vol. 15, No. 6, 2001, pp. 30-39.

[15] J. Wu, "Optimal broadcasting in injured hypercubes using directed safety levels," *Journal of Parallel and Distributed Computing*, Vol. 63, No. 9, 2003, pp. 815-826.

[16] ——, "Fault-tolerant adaptive and minimal routing in mesh-connected multicomputers using extended safety levels," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 11, No. 2, 2000, pp. 149-159.

[17] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice, Fourth Edition*. Springer-Verlag, 1997.

[18] A. Jadbabaie, "On geographic routing without location information," *Proc. of the* 43^{th} *IEEE Conference on Decision and Control*, Dec. 2004, pp. 4764-4769.

[19] M. Sichitiu and V. Ramadurai, "Localization of wireless sensor networks with a mobile beacon," *Proc. of the* 1^{st} *IEEE International Conference on Mobile Ad Hoc and Sensor Systems (IEEE MASS'04)*, Oct. 2004, pp. 174-183.

[20] S. Capkun, M. Hamdi, and J. Hubaux, "GPS-free positioning in mobile ad hoc networks," *Proc. of the* 34^{th} *Annual Hawaii International Conference on System Sciences*, 2001, pp. 3481-3490.

[21] Q. Fang, J. Gao, and L. Guibas, "Locating and bypassing routing holes in sensor networks," *Proc. of the* 23^{rd} *Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM'04)*, 2004, pp. 2458-2468.

[22] S. Chen, G. Fan, and J. Cui, "Avoid "void" in geographic routing for data aggregation in sensor networks," *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 1, No. 4, 2006, pp. 169-178.

[23] T. Dimitriou and I. Krontiris, "GRAViTy: Geographic routing around voids in sensor networks," *International Journal of Pervasive Computing and Communications*, Vol. 2, No. 4, 2006, pp. 351-361.

[24] R. Kleinberg, "Geographic routing using hyperbolic space," *Proc. of the* 26^{th} *Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM'07)*, 2007.

[25] N. Arad and Y. Shavitt, "Minimizing recovery state in geographic ad-hoc routing," *Proc. of the* 7^{th} *ACM International Sysmposium on Mobile Ad Hoc Networking and Computing (ACM MobiHoc'06)*, May 2006.

[26] D. Wang, "A rectilinear-monotone polygonal fault block model for fault-tolerant minimal routing in meshes," *IEEE Transactions on Computer*, Vol. 52, No. 3, March 2003, pp. 310-320.