Optimization of Data Analysis Algorithms for Geographic Information System

Xiaoguang An

Management Engineering Department, Hebei Vocational College of Rail Transportation, Shijiazhuang 050051, China

Corresponding Author Email: anlistone@163.com

ABSTRACT

Conventional network protocols largely rely on global information and their scalability is usually unsatisfactory. This study performed simulation experiments on conventional network protocol algorithms based on network topology and attained a conclusion that the data packet delivery performance and overhead performance of proactive routing algorithms is better than that of reactive routing algorithms. By summarizing existing studies and analyzing the routing decision of greedy algorithms, it's found that the greedy routing algorithm based on closest distance criterion can avoid generating loops so it was selected the basis of local routing decisions; besides, the rate of successful data packet delivery and the average path length were taken as performance indicators, and the results of simulation experiments showed that the unilateral traversal face routing method has a shorter path and a better performance. This study proposed a new method that integrates the merits of the greedy routing forwarding method and the face routing forwarding method and achieved very good data forwarding performance. Based on the local link quality, a correction parameter was introduced to optimize the routing protocol. Then simulation was performed on the algorithm before and after optimization, and the results proved that the optimized routing protocol has higher data packet delivery rate, smaller data packet overhead, better scalability, and higher data forwarding performance.

1. INTRODUCTION

Thanks to the fast development of computer and wireless transmission technologies, information exchange and transmission have becoming increasingly convenient these days [1]. Now the one-way communication can no longer meet user needs in real-time data sharing at low costs, and the construction of networks not only consumes large amounts of material resources such as electric and optical cables, but also requires high maintenance costs [2].

Wireless ad hoc networks are a type of networks consisting of nodes connected through wireless links, since no physical circuit is needed, their configuration is flexible and the network expansion is easier, now they are a hot topic in information research [3]. Algorithm selection can affect the form and effect of ad hoc networks, so algorithm research is particularly important [4]. Conventional routing algorithms generally attain the shortest path through weighting. The traditional approaches usually require to build routes, and the process is complex, time-consuming, and costly [5]. In contrast, greedy algorithms make local decisions based on local data, there is no need to establish an entire route. On the one hand, it can reduce the maintenance cost of the network, on the other hand, the implementation effect of the algorithm can be guaranteed in case of high data volume, and the delay problem won' t show up [6].

In view of these matters, this study aims at studying and optimizing the greedy routing algorithm based on geographic location information, in the hopes of attaining higher performance and better serving the user needs.

2. BASIS OF ROUTING ALGORITHMS

For wireless ad hoc networks, the routing problem is equivalent to the problem of finding the shortest path. The topology-based routing algorithms are a common approach for solving the routing problem, they can be divided into proactive algorithms and reactive algorithms [7], and their respective characteristics and expressions are given in Table 1. The basic unit of a routing algorithm is the routing table, nodes in the routing table are presented as neighbour nodes, an ID number is set for each data packet. The data packets are forwarded continuously, the data nodes and data identifiers are transmitted and traversed, until reaching the destination node.

Table 1. Common routing algorithms

Routing Algorithm	Characteristic	Typical Routing Protocol Algorithm
Proactive routing protocols	1. Need to maintain routing tables. 2. With smaller time delay. 3. Routing table maintenance needs to consume large amounts of network bandwidth resources of energy and power.	DSDV protocol (The most common), WRP protocol, OLSR protocol, STAR protocol, TBRPF protocol
Reactive routing protocols	1. Do not need to periodically maintain routing tables. 2. Effectively reduce the network bandwidth and the node energy consumption.	DSR protocol, AODV protocol, ABR protocol, TORA protocol

2.1 Performance analysis of conventional routing protocol algorithms

The transmission performance of conventional routing protocol algorithms differs according to scenarios [8]. This paper determined two kinds of performance of conventional routing protocol algorithms via simulation, one is the data packet delivery performance, the other is the data packet overhead performance. Parameters of the simulation experiment are show in Table 2. A 1500m×300m rectangular space was selected, and the parameters of each node in the wireless network are shown in Table 3 (similar to the parameters of Lucent Direct Sequence Spread Spectrum WLAN card).

Table 2. Experimental simulation parameters

Experimental Simulation Parameters	Parameter Values
Node number	50
Area length	1500m
Area width	300m
simulation time	900s
Residence time node movement	0s, 30s, 60s, 120s, 300s, 600s, 900s
The largest mobile node rate	20m/s
Data flow type	CBR
Data packet size	64Byte
Data flow rate	4p/s
Data flow logarithmic	30

Table 3. Wireless network card parameters

Wireless Network Card Parameters	Wireless Network Card Parameter Values
signal sensing threshold	10
Carrier to listen to the threshold	15.59pW
receiving signal threshold	365.2pW
Launch the bandwidth	2MHz
Transmission power	0.28W
wireless signal frequency	915 MHz

There are two evaluation indicators:

(1) Data packet delivery performance. It indicates the rate of successful conversion and transmission of data packets, and its expression is as follows:

Data packet delivery rate = Number of data received by the destination node / Number of data packets sent by the source node at a fixed rate;

(2) Data packet overhead performance. It indicates the sum of routing protocol packets during the simulation analysis.

Changes in the two indicators with the time of node movement are shown in Figures 1 and 2. As can be observed from the figures, when the retention time of node movement was shorter, the delivery performance of DSR routing algorithm and AODV routing algorithm was good; when the retention time of node movement was longer, namely the moving speed was lower, the delivery performance of the three algorithms was high, that is, the successful rate was good. According to Figure 2, when the retention time of node movement was shorter, namely the movement was intense, the data packet overhead performance of the AODV routing algorithm was high; while for the DSDV routing algorithm, its overhead performance didn't change with the moving speed of node.

Above research suggested that, the data packet overhead

performance of the DSDV routing algorithm is good, and its data packet delivery performance is good as well, so it had been chosen as the theoretical basis of this paper to carry out subsequent research.
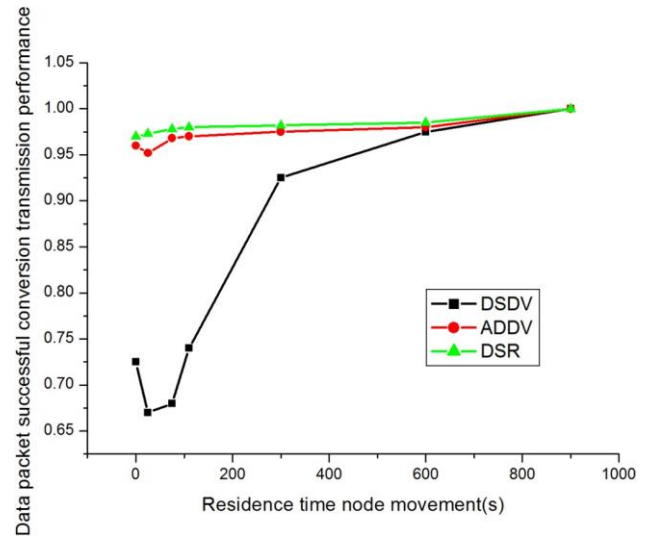

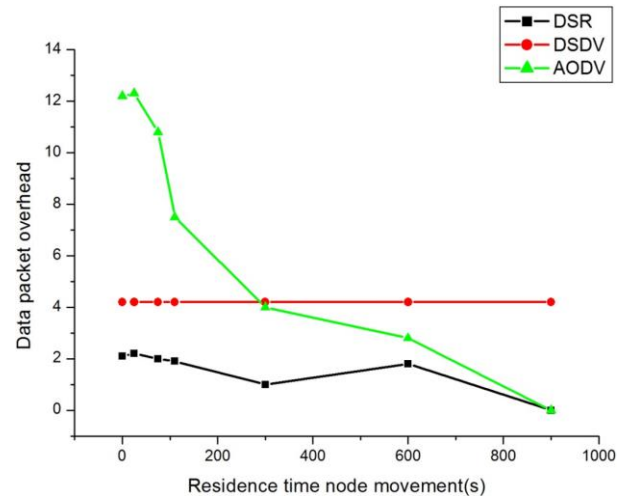**Figure 1.** Data packet delivery performance

Figure 2. Data packet overhead performance

3. ABOUT THE NEW NETWORK GREEDY ALGORITHM

3.1 Routing decision analysis

Greedy algorithm focuses on the current optimal decision as in some cases a locally optimal is also the global optimal. The main idea of greedy decision-making algorithm is to make a series of immediate optimal decisions, then as the decision range gradually narrows down to the range of all decisions, the global optimal decision could be found [9]. The greedy routing algorithm based on the information of geographical location is different from other algorithms as its data transmission method adopts a hopping approach to get closer to the destination node with its address as the basis. Its data storage doesn't need the help of routing tables or network addresses, for dynamic networks, it can reduce additional data packet transmission, so continuous upgrading and maintenance is not necessary, and the broadband usage can be reduced, therefore, it is quite

competitive. According to the similarities and differences in the judgement criteria of the algorithm, common greedy routing algorithms can be divided into three types: the greedy routing algorithm based on direction recently criterion, the greedy routing algorithm based on prior to the recent criterion, and the greedy routing algorithm based on closest distance criterion. The criteria and characteristics of these three algorithms are summarized in Table 4.

Table 4. Algorithm classification

| Algorithm	Judgement Criterion	Characteristics
Direction recently criterion of greedy routing algorithms	The current node selection closest to the direction of the destination node neighbor node as the next-hop node.	For network nodes without memory, may cause a routing loop, waste a lot of storage space.
Prior to the recent criterion of greedy routing algorithm	Prior to the shortest distance.	Circuit will not occur.
Closest distance criterion of greedy routing algorithm	Based on the physical location of the node information, Physical location nearest.	Circuit will not occur Apply to three-dimensional space.
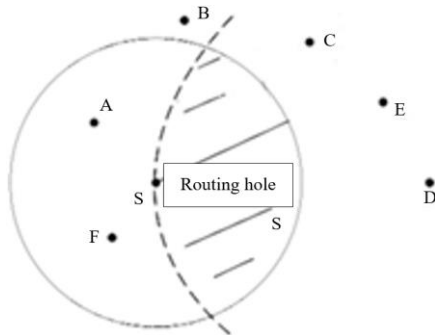
Figure 3. The routing hole problem

The three algorithms have one thing in common: they select the nearest node for data forwarding according to their respective judgement criteria based on current node, when the current node doesn't match with the target node, namely the next receiving node can not be found, then as shown in Figure 3, the neighbor nodes corresponding to current node S are B, C, E, and D, these neighbor nodes are not "closer" nodes for node S, then the routing hole problem would arise. Reasons for the first case of routing hole problem include obstacles, unreliable nodes, etc. Data will be lost at this time, which is not conducive to the implementation of the algorithm.

To further figure out the performance of each algorithm, based on the ad hoc network, this paper referred to the two-factor theory (node number n and node connection number m) and comparatively analyzed the performance of these three algorithms in the ad hoc network, the average packet delivery rate and hop number were selected as performance parameters. The selected experimental parameters are shown in Table 5. After averaging the results of 20 network connectivity graphs, the relationships between the average packet delivery rate and hop number are given by the curves shown in Figures 4 and 5.

Table 5. Experimental parameters

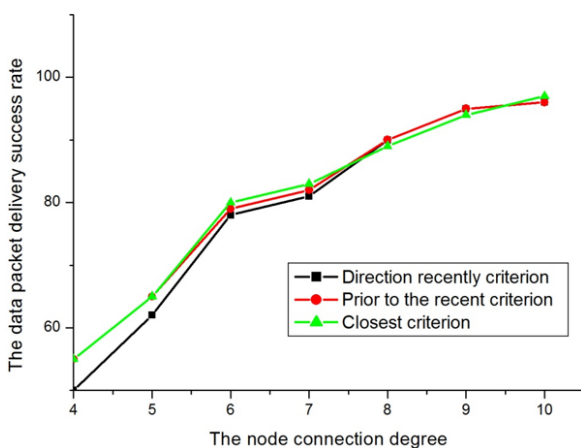|	Values	Note
Node number	20,50,100	0-100
Network connectivity graph	20	Nodes in the wireless transmission radius is sorted first nm / 2 in all side length
Source node and destination node	50	

Figure 4. The data packet delivery rate

Figure 5. The path hop on average

According to above results, the greedy routing algorithm based on closest distance criterion and the greedy routing algorithm based on prior to the recent criterion outperformed the greedy routing algorithm based on direction recently criterion, in this paper, the greedy routing algorithm based on closest distance criterion was chosen as the basis of routing decisions.
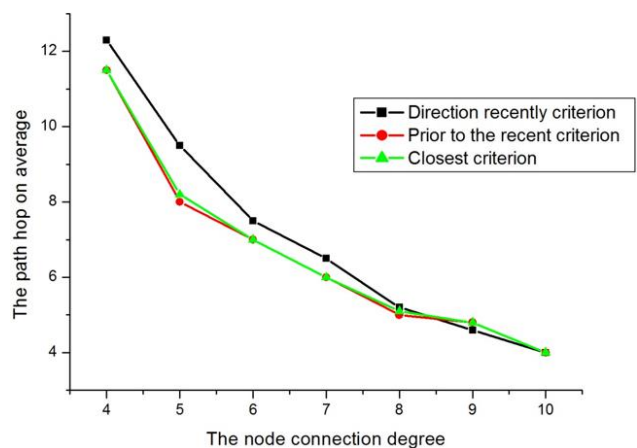
3.2 Processing of the greedy routing algorithm

The routing hole problem is a challenge that can be avoided by using a greedy routing algorithm based on the information

of geographical location to collect local information as much as possible so as to reduce the use of network information and the amount of data flow. At the position of each node, the greedy routing algorithm uses the local position information of neighbour nodes to create a Related Neighbour Graph (RNG), as pointed out by Giordan and Stojmenovic [10], RNG has the same connectivity as the GG (Figure 6), and the expression is:

$$\forall w \neq u, v: d(u,v) \leq max[d(u,w), d(v,w)]$$

where,
u/v/w are independent vertices;
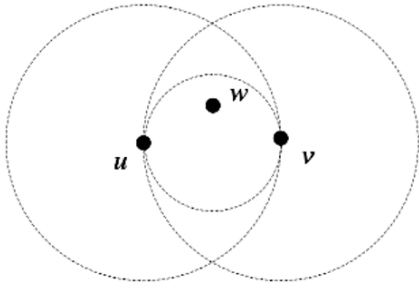$d(u,v)$, $d(u,w)$, $d(v,w)$ are distances between nodes u and v, u and w, and w and v.

Figure 6. RNG and GG

The nodes in RNG that do not meet the conditions of the formula were filtered out by the following algorithm:
m is the midpoint of edge $\overline{uv}$
for all $v \in N$ do
for all $w \in N$ do
if $w==v$ then
continue
else if $d(m,w)<d(u,m)$ then
delete edge $\overline{uv}$
break
end if
end if for
end if for

3.3 Analysis and optimization of greedy routing algorithm

3.3.1 Analysis of greedy routing algorithm
The process of data packet routing and forwarding is the complanation process of the connectivity graph. There are two methods of data forwarding: the complete traversal and the unilateral traversal methods [11]. To figure out the performance of face routing data forwarding, the paper adopted simulation experiments to study the rate of successful data delivery and the average path length extension. The selection criteria of simulation parameters are similar to those given in Table 5 of Section 3.1, and the performance parameters are:
(1) Data packet delivery rate:

$$DR_A(G) = \frac{|X|}{n(n-1)}$$

where,
n represents the number of nodes;
A represents the algorithm;
X represents the set of node pairs $(u,v) \in G: u \neq v$;
$|x|$ represents the number of elements in set X.

(2) Average path length extension:

$$AD_A(G) = \frac{1}{|X|} \sum_{(u,v)} \in X \frac{AP(u,v)}{SP(u,v)}$$

where,
$AP(u,v)$ represents the number of edges in the network graph through which the effective path from node u to v passes;
$SP(u,v)$ represents the number of edges in the network graph through which the shortest path from node u to v passes (Dijstra's algorithm).
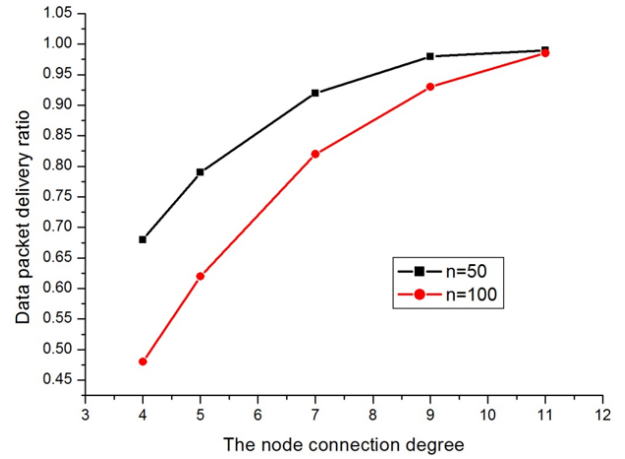

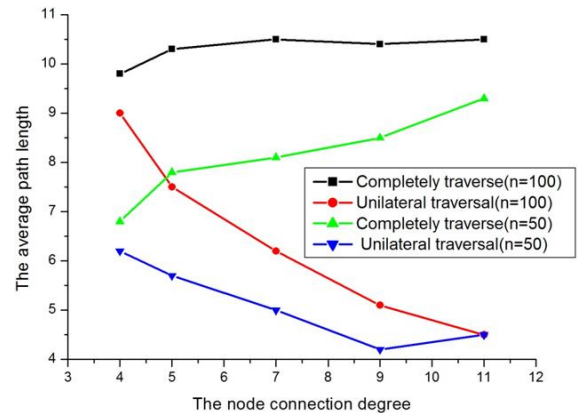**Figure 7.** Performance of greedy routing algorithm

Figure 8. Average path length of face routing and forwarding method

According to Figures 7 and 8, in terms of average path length extension, the performance of unilateral traversal is better than that of complete traversal. After the route forwarding method and the greedy algorithm were combined, the average path length extension was good, the data packet delivery rate had been improved, the merits of the two had been combined as well.

3.3.2 Optimization of greedy routing algorithm
The greedy routing algorithm protocol is a good reference for solving the routing hole problem during the forwarding process of data packets. However, when two receiving nodes have a same length value relative to the destination node and different distances relative to the source node, the data forwarding delivery of the algorithm is prone to failure [12]. In view of this problem, this paper introduced a correction

parameter *f* to describe the quality influence of such local link, and the expression of *f* is:

$$\begin{cases} f = tanA(d - B) + C \\ \quad d \rightarrow d_c : f \rightarrow 1 \\ \quad d \rightarrow R_{max} : f \rightarrow \infty \end{cases}$$

Table 6. Simulation scenarios

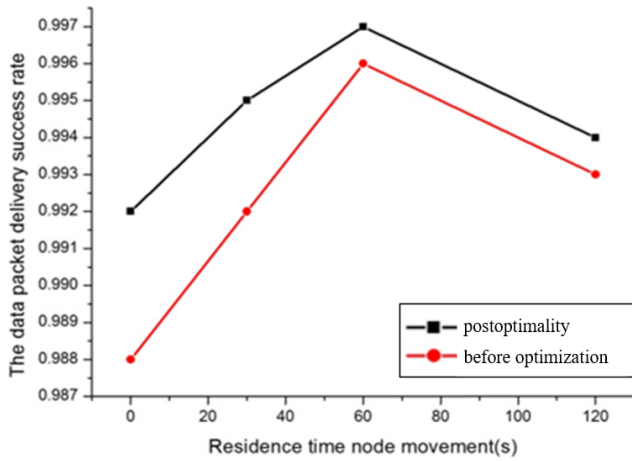| Number of Nodes	Simulation Area	Node Density	CBR Stream
48	1500m×300m	$1node/9000m^2$	20
100	2000m×400m	$1node/9000m^2$	20

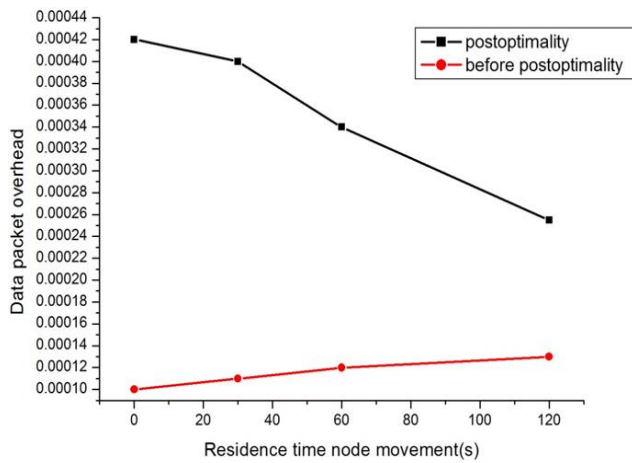**Figure 9.** Rate of successful data packet delivery
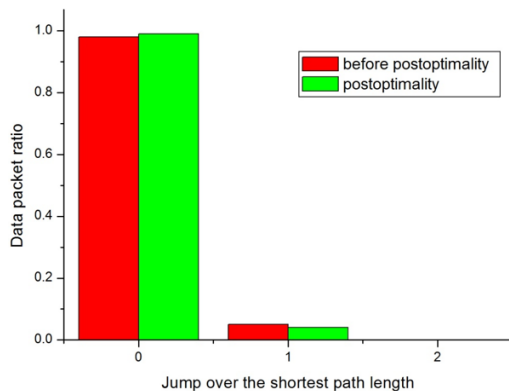

Figure 10. Data packet overhead

Figure 11. Data packet ratio

To verify the reasonableness and correctness of the optimization method, simulation experiment was carried out for analysis, and the experimental parameters are shown in Table 6, for other experimental parameters, please see Section 3.1. Three indicators, data packet delivery rate, path length, and algorithm overhead were selected for analysis, then the attained three parameters and node motion situations are shown in Figures 9-11.

According to the three figures above, as the retention time of node movement changes, the data packet delivery rate of the optimized greedy routing algorithm protocol is higher than that before optimization, the corresponding protocol routing has a less overhead, while for the average path length of the data algorithm, there is not a large difference before and after optimization. Overall speaking, the quality optimization of local link based on correction parameter can improve the rate of successful data packet delivery.

4. CONCLUSION

This paper analyzed the conventional routing protocol algorithms and greedy algorithms, carried out simulation experiments to verify their performance, introduced a correction parameter to solve the routing hole problem, and attained several conclusions:

(1) Conventional routing protocol algorithms can be divided proactive-style and reactive-style routing algorithms, through simulations, it's verified that the data packet delivery rate and overhead of proactive routing algorithm are better than those of the reactive routing algorithm.

(2) Greedy algorithms and routing decisions were analyzed and the simulation results showed that the greedy routing algorithm based on closest distance criterion can solve the routing loop problem to some extent and its performance is better than that of the greedy routing algorithms based on prior to the recent criterion and based on direction recently criterion.

(3) With data packet delivery rate and average path length taken as performance indicators, the performance of complete traversal and unilateral traversal methods was analyzed through simulation, and it's verified that the unilateral traversal outperformed the complete traversal in terms of the two performance indicators.

(4) Based on local link quality, a correction parameter was introduced to optimize the routing protocol. The performance of the algorithm before and after optimization was simulated experimentally, and the results showed that the optimized algorithm has a higher data packet delivery rate and a smaller overhead.

REFERENCES

[1] Li, X.H., Hong, S.H., Fang, K.L. (2011). WSNHA-GAHR: A greedy and A* heuristic routing algorithm for wireless sensor networks in home automation. IET Communications, 5(13): 1797-1805. https://doi.org/10.1049/iet-com.2010.0746

[2] Mitzenmacher, M. (1996). Bounds on the greedy routing algorithm for array networks. Journal of Computer and System Sciences, 53(3): 317-327. https://doi.org/10.1006/jcss.1996.0072

[3] Ryu, M.W., Cha, S.H., Koh, J.G., Kang, S., Cho, K.H. (2011). Position-based routing algorithm for improving

reliability of inter-vehicle communication. KSII Transactions on Internet & Information Systems, 5(8): 1388-1403. https://doi.org/10.3837/tiis.2011.08.002

[4] Mohamed, A.A., Sheik, A.K.P., Munir, A.R.M. (2011). Gateway node-based greedy routing algorithm for efficient packet forwarding in vehicular ad hoc networks. International Journal of Engineering Science & Technology, 3(5): 3640-3650. https://doi.org/10.3850/978-981-07-3043-7_030

[5] Sreenivasarao, V., Potluri, C.S., Kadiyala, S., Yohannes, G. (2012). Improving performance analysis of routing efficiency in wireless sensor networks using greedy algorithm. International Journal of Advanced Research in Artificial Intelligence, 1(2): 45-49. https://doi.org/10.14569/ijarai.2012.010208

[6] Li, X.H., Hong, S.H., Fang, K.L. (2011). Location-based self-adaptive routing algorithm for wireless sensor networks in home automation. Eurasip Journal on Embedded Systems, 2011: 484690. https://doi.org/10.1155/2011/484690

[7] Liu, W.J., Feng, K.T. (2009). Greedy routing with anti-void traversal for wireless sensor networks. IEEE Transactions on Mobile Computing, 8(7): 910-922. https://doi.org/10.1109/tmc.2008.162

[8] Heissenbüttel, M., Braun, T., Bernoulli, T., Wälchli, M. (2004). Blr: Beacon-less routing algorithm for mobile ad hoc networks. Computer Communications, 27(11): 1076-1086. https://doi.org/10.1016/j.comcom.2004.01.012

[9] Wu, Q.Y., Liu, X.H., Jing, N. (2011). Coverage-preserving multi-path routing algorithm for wireless sensor networks. Advanced Materials Research, 225-226: 199-202. https://doi.org/10.4028/www.scientific.net/amr.225-226.199

[10] Giordano, S., Stojmenovic, I. (2004). Position based routing algorithms for ad hoc networks: A taxonomy. Network Theory & Applications, 14: 103-136. https://doi.org/10.1007/978-1-4613-0223-0_4

[11] Abdallah, A.E., Fevens, T., Opatrny, J., Stojmenovic, I. (2010). Power-aware semi-beaconless 3D georouting algorithms using adjustable transmission ranges for wireless ad hoc and sensor networks. Ad Hoc Networks, 8(1): 15-29. https://doi.org/10.1016/j.adhoc.2009.03.001

[12] Chun, J., Shioura, A., Tien, T.M., Tokuyama, T. (2014). A unified view to greedy geometric routing algorithms in ad hoc networks. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 97(6): 1220-1230. https://doi.org/10.1587/transfun.E97.A.1220