2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)

A Novel P2P Service Discovery Algorithm Based on Markov in Internet of Things

Bing Jia
School of Computer Science
Inner Mongolia University, 010021
Hohhot, China
jiabing@imu.edu.cn

Wuyungerile Li
School of Computer Science
Inner Mongolia University, 010021
Hohhot, China
gerile@imu.edu.cn

Tao Zhou
School of Computer Science
Inner Mongolia University, 010021
Hohhot, China
1656605330@qq.com

***Abstract*—In order to adapt to the characteristics of high mobility, low density in Internet of Things, this paper proposes a novel P2P service discovery algorithm based on Markov. The Markov is used to achieve the prediction of the meeting time interval between any two nodes, and to construct a two-stage prediction model of meeting time interval. This paper uses the model to improve the Spay and Wait routing protocol to improve the delivery success rate of service message. The experiments show that the algorithm can improve both of the recall rate and precision without sacrificing too much time.**

Keywords-Spray and Wait; Markov; P2P service discovery; Internet of Things

I. Application Scenarios for P2P Service Discovery Algorithm

In the Internet of Things (IoT), the mobile AdHoc network and vehicular networks are very important as the representative of the new network. The new network is especially suitable for the field or the open environment. And a mobile user needs to obtain the service information in the area. For example, when a fireman is dealing with the wildfires, he may need to ask for a periodic update of the temperature map. The range of the map is within a kilometer to the fire on the surrounding in its position. This map will keep him on the alert better. In the scene, the fireman can be seen as a mobile node. And the temperature sensor also possibly has the mobility through the vehicle [1]. Another example, when a mobile robot is searching and rescuing in the unknown dynamic environment, it may need to check the sensor information on terrain and survivors about the surrounding environment. Based on the query results, the robot is able to locate survivors and find the best rescue route through the motion planning. In this case, the original service discovery mechanism via a fixed centric registration will not be able to work, because it is impossible to require all service provider nodes to register in a so called UDDI by the downed communication networks, and also carry out the service discovery in the space. It needs to set up another kind of discovery algorithm for searching and matching service node in service area. In order to obtain the required services, it needs to use P2P service discovery algorithm. Fig. 1 shows the process of P2P service discovery.

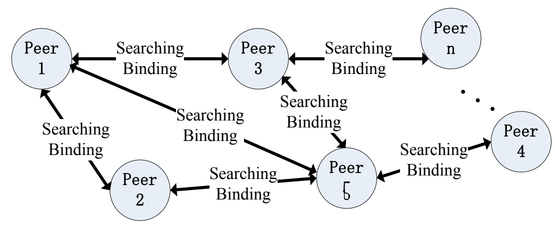

Figure 1. P2P service discovery

As shown in Fig. 1, the user's role can be "service user ", and also can be "service provider" in P2P framework. The number of "service provider" consciously changed with the number of "service user". The more people used, the more people served. The mode is flexible, which service nodes can join and leave at any time, without affecting the structure of the service network. And it is easy to expand. Although in the condition that the cluster head and the cluster node may have formed a structured service organization, there is still a P2P between the cluster heads. And the process of service discovery is mainly carried out between the cluster heads. So it needs to use a P2P service discovery. It is different with the complete P2P structure service discovery algorithms. Its searching object is the service catalog of the whole cluster which is located in the cluster head.

II. Related Works

In traditional P2P service discovery algorithm, it usually uses multicast or broadcast manner to find the services met the requirement of the service requesters [2-11]. Most of the service requesters use flood policy for distributing their requests on the network, or selectively forward based on the cache directory information of service nodes. The cost of the establishment and maintenance of the service cache list may be large, and the energy of the node consumed by the global broadcast may be great, which can't meet the principles of IoT services design. In addition, these algorithms are more suitable for the ad hoc network which has a relatively stable link. And these algorithms cannot be suitable for the Delay Tolerant Networks (DTN) caused by dramatically bad communication environment, since the node's high mobility and low density, and other factors. In order to meet the service discovery in DTN, we need a DTN routing algorithm equipped with service information to implement the service request message and response message forwarding. Finally it can achieve the purpose of service discovery.

In the study of routing in DTN, Epidemic [12] is the basic routing algorithm based on copy, but the large sum of copies will consume a lot of network resource which lead to

978-1-5386-3221-5/17 $31.00 © 2017 European Union
DOI 10.1109/CSE-EUC.2017.191

serious network congestion. Many scholars have proposed many methods to improve Epidemic, the widely used is Spray and Wait [13], which is a multi-copy routing algorithm. It uses the copy less than the amount of Epidemic, and achieves approximate performance of Epidemic. In this paper, we use Markov model to predict the meeting time interval between the nodes, and add the predict result to the Spray and Wait routing algorithm. Then, find out the node that recently met with the destination node as the highest utility value node for service message delivery. In this paper the service discovery algorithm based on Markov meeting time interval prediction function model is called FMSD for short.

III. MARKOV-PM

It is not accidental that the nodes are encounter in some scenarios, and there is an inherent law. Based on the interval sequence of nodes in the previous encounter, we use Markov chains to predict the approximate range for the next encounter interval, in order to find out a node that may have first encounter with the destination node to forward the message. In this paper, the approach is called a Markov time interval prediction model (Markov-PM).

We supposed the meeting interval as a random variable X. The sequence X_i can meet the requirement of discrete time and discrete state by Markov chain [14]. For any two nodes in the network, they have the whole interval sequence X_i about their encounter, and the X_i in accordance with given range is divided into seven states. When $0 < X_i \le 100$ the state is "1"; when $100 < X_i \le 200$ the state is "2"; when $200 < X_i \le 300$ the state is "3" when $300 < X_i \le 400$ the state is "4"; when $400 < X_i \le 600$ the state is "5"; when $600 < X_i \le 800$ the state is "6"; when $800 < X_i$ the state is "7". Therefore, the encounter between any two nodes can be represented as sequence a_i that is consisted by the state "1-7". The properties of Markov chain is as follow:

$$P(X_n{=}a_n | X_{n-1}{=}a_{n-1}, X_{n-2}{=}a_{n-2}, \ldots, X_1{=}a_1) = P(X_n{=}a_n | X_{n-1}{=}a_{n-1}) \quad (1)$$

Wherein, X_i denotes the interval of each twice encounter, a_i shows a state which the encounter time interval belongs to, ranging from "1" to "7". X_i is counted by each node itself, and stored in the local array in the form of a_i. When a source node is encountering with any node B_i, the first thing is to update their encounter time interval sequence. And then it will view the historical encounter interval sequence of B_i which have the encounter with the destination node and establish the state transition matrix "P" by the historical sequence, as shown in Eq. (2).

$$p = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} & p_{17} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} & p_{27} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} & p_{37} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} & p_{47} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} & p_{57} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} & p_{67} \\ p_{71} & p_{72} & p_{73} & p_{74} & p_{75} & p_{76} & p_{77} \end{pmatrix} \end{array} \quad (2)$$

Transfer matrix P_{ij} is the probability of the transfer from the state i to state j.

$$\sum_{i=1}^{7} P_{ki} = 1, (k = 1,2,3,4,5,6,7) \quad (3)$$

Assuming the current state is k, it will find out the column corresponding to the maximum number of probability value in row k in the state transition matrix. So it can predict the meeting interval that the node will meet the destination node.

Suppose the encounter time interval sequence that a node S_a have an encounter with the destination node S_b is {2, 1, 3, 2, 4, 5, 4, 1, 3, 2, 3, 6, 7, 5} The state transition matrix can be obtained as (4):

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 0 & 1/3 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \end{array} \quad (4)$$

Since the last state of S_a is "5", by the state transition matrix (4), we can look for all the probability values in the fifth row. The maximum probability value "1" is in the fourth column, so we can forecast that the next interval state is "4", which can predict that the node S_a will have an encounter with the destination node S_b in the interval time (300,400].

According to Markov-PM, we predict the state of the next encounter interval a_i for each B_i node which will have an encounter with the destination service node. The smaller the a_i, the earlier encounter with the destination service node. However, there may exist more than two B_i that will have an encounter with the destination service node within the same time interval a_i, in this paper, we adopt two-step prediction methods, which used the same state transition matrix to predicted a step further. If there are still more than two B_i that will have an encounter with the destination service node within the same time interval a_i after two-step prediction, we will compare the time that have an encountering with the destination node between these nodes, and select the node whose encounter time with the destination service node is earliest as the message delivery node. Fig.2 shows the forwarding process of Markov-PM.

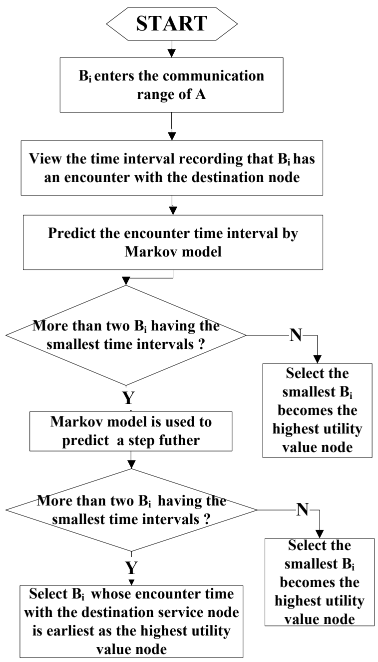

Figure 2. The forwarding process of Markov-PM

IV. SPRAY AND WAIT BASED ON MARKOV-PM

Spray and Wait [13] is a classic routing protocol that has a fixed quota copy. The message transmission process is divided into "spray stage" and "wait stage". First, we determine the number of the message copies in the source node. When the node has an encounter with the other nodes, if the node does not have the message, it will spray a portion of the message copies to it. Until the node only has one message copy, it will spray no longer and enter the "wait" state. It will send the only message copy to the destination node until it has an encounter with the destination node. After that, the node that has received the message copy will also be performed like this. The algorithm can ensure that the number of the message copies in the entire network within a certain range to reduce the consumption of the network resource

Both the two stages of Spray and Wait are obvious deficiencies. During the "Spray stage", since it selects a node to distribute randomly, there may be a large number of messages may be transmitted to the low utility value node. During the "Wait stage", since it waits for the destination node passively, it may miss many high probabilities of efficient nodes that will have an encounter with the destination nodes. The Markov-PM is applied in both two stages of Spray and Wait. It can improve the above two issues, as shown in Fig. 3. After the Markov-PM is applied in the "Spray stage", the efficient nodes that will have an encounter with the destination node earlier within the communication radius may be found to distribute the message. Each node uses the same spray method to do a binary distribution until it has only one message copy, then it enters to "Wait stage". It will deliver the unique message to the node that may meet the destination node earlier based on Markov-PM, rather than waiting for the destination node passively.

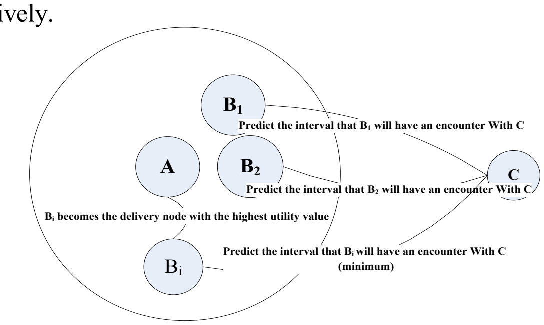

Figure 3. The prediction of encounter

V. THE INTERACTIVE PROCESS

Fig. 4 describes the interaction process when a client sends a service request. The main steps are as follows:

1) Client access: the client selects a node in the ad hoc network as an access node, which is a common node. Only special feature is that it has the communication with connected equipment by the automatic protocol conversion. In this case, the access node can notify other nodes in the area to turn on the power, and wait for interaction [15].

2) Forward the service request: the client forwards the service request package to the communicably-capable service node via the access node. The access node will send the searching request to other nodes of the network according to routing policy.

3) Receive the service request: the nodes of the network will match the service provided by it after receiving the request. If the matching is successful, it will return the response to the access point. The service response includes a service node provided a detailed description of the document.

4) Use the service: Once the client receives a service response, it knows the capabilities of this node and the possible actions. The client can send control information to the node requesting to use the service. And the node will return data or service to the client.

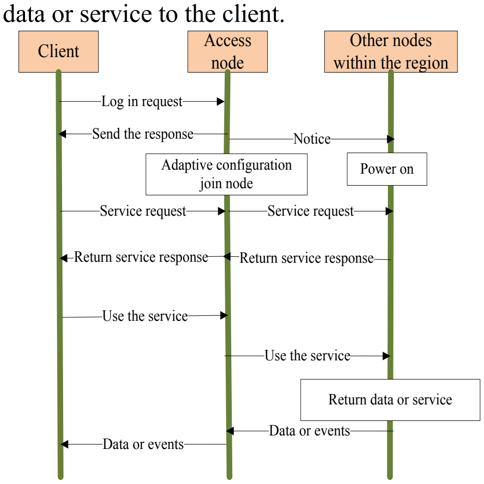

Figure 4. The interactive process of service requests

VI. SIMULATION AND ANALYSIS

We use the simulator ONE (Opportunistic network environment simulator) to simulate the proposed FMSD service discovery algorithm. We use the default map of ONE: the map of the Helsinki city. We set the number of the service nodes for 50, 100, 150, 200 and 250. No matter what kinds of the amount of the service nodes, there are 25 kinds of services. That is, if the number of the service nodes is 100, the service is divided into 25 types. Each type of the service are four service nodes. The simulation parameters are set as follow:

TABLE I. THE SIMULATION PARAMETERS

Parameter	Value
Simulation time	18000s
Simulation area size	4500m*4500m
Number of nodes	50，100，150，200，250
Mobility rate	3-4m/s
pace time	(1-120)s
Transmission rate	250Kbps
Transmission range	50m
Cache size	40M
Message generation interval	[300，400] [400，500] [500，600] [600，700]
TTL	3600s

In the above simulation scenario, we compare the FMSD algorithm with the service discovery algorithm based on Epidemic (Epidemic-SD), and the service discovery algorithm based on Spray and Wait (Spray and Wait-SD). The average recall, the average precision, the average delay both of the service request and response of the three algorithms are compared.

When the node message generation interval (MGI) is 500s-600s, and the initial number of copies is 6. The total number of nodes is set as 50,100,150,200 and 250 in network. Fig. 5 and Fig. 6 are given the average recall and average precision of the three algorithms for comparison. Fig. 7 and Fig. 8 show the average delay of the service request and response for comparison.

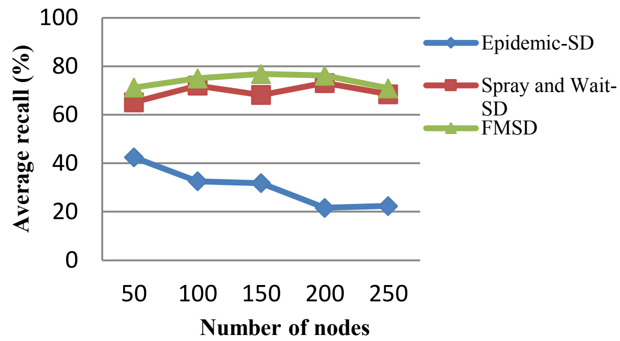

Figure 5. Average recall for different number of nodes

As shown in Fig. 5, the average recall of the proposed algorithm FMSD is better than the other two for any kind of the amount of the service nodes.

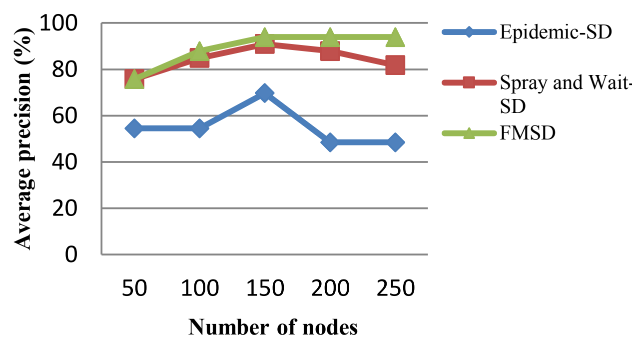

Figure 6. Average precision for different number of nodes

As shown in Fig. 6, the lowest average precision is Epidemic-SD. The improvement of the average precision of FMSD is not obvious in the 50 service nodes comparing with Spray and Wait-SD. But with the increase of the amount of the service nodes, FMSD shows good performance. When the number of nodes is 250, it is higher than Spray and wait-SD nearly 10%.

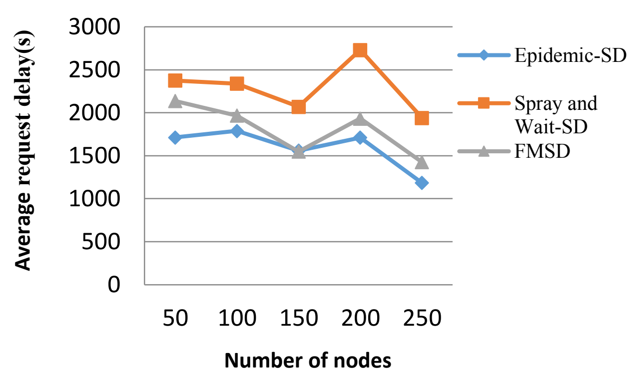

Figure 7. Average request delay for different number of nodes

As shown in Fig. 7, the average request delay of FMSD is less than Spray and Wait-SD for different number of nodes. Comparing with Epidemic-SD, only when the number of nodes is 150, it is less than Epidemic-SD. because Epidemic-SD will abandon the longer delay message automatically.

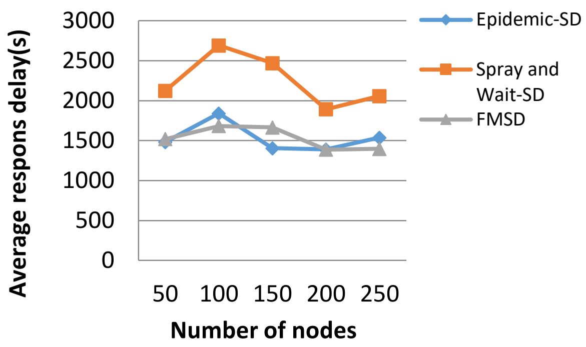

Figure 8. Average respons delay for different number of nodes

As shown in Fig. 8, the service response average delay of FMSD is significantly less than Spray and Wait-SD, and is basically consistent with Epidemic-SD. In addition, we find that the average response delay of FMSD is less than the average request delay, because the application of Markov-PM may be consuming some time.

In order to test the impact of MGI in service discovery, we complete another experiment. The number of nodes is set for 100, the initial number of copies is 6, the same proportion of service types. Message generation interval is adjusted in four cases, the average recall and precision of the three algorithms are shown in Fig. 9 and Fig. 10, the average request delay and the average response delay are shown in Fig. 11 and Fig. 12.

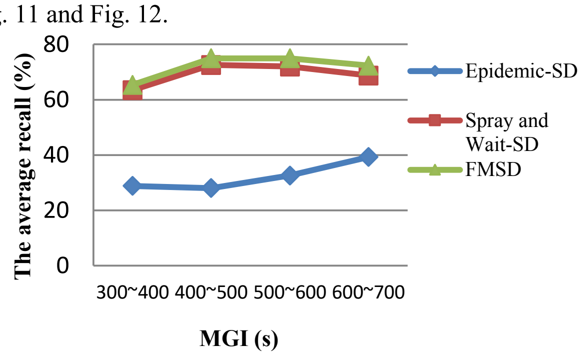

Figure 9. Average recall for different MGI

As shown in Fig.9, the average recall of FMSD is the highest for the four kinds of MGI. The average recall of FMSD is more than Spray and Wait-SD algorithm average nearly 5 percentage points, and more than Epidemic-SD algorithm nearly 35 percent. The MGI has little effect on the FMSD algorithm.

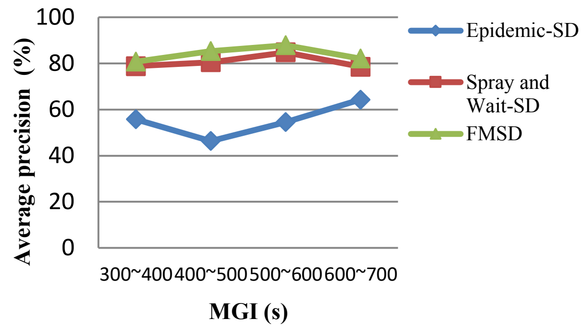

Figure 10. Average precision for different MGI

As shown in Fig.10, the average precision of FMSD is more than Spray and Wait-SD for the four kinds of MGI. Comparing with Epidemic-SD, it has greatly improved. And the average precision of FMSD reaches almost 90% when MGI is 500s-600s.

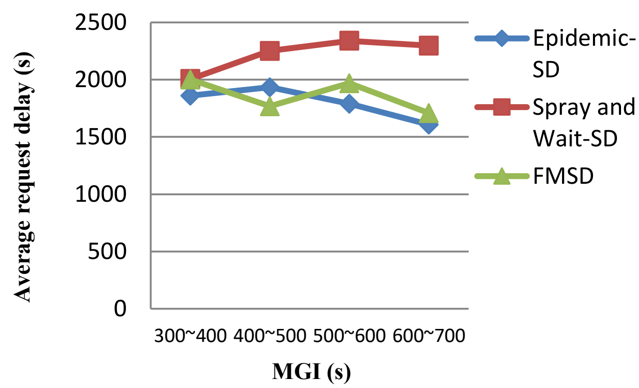

Figure 11. Average request delay for different MGI

As shown in Fig.11, FMSD obtained the minimum average request delay when MGI is 600s-700s. The average response delay of FMSD is less than Spray and Wait-SD and substantially consistent with Epidemic-SD.

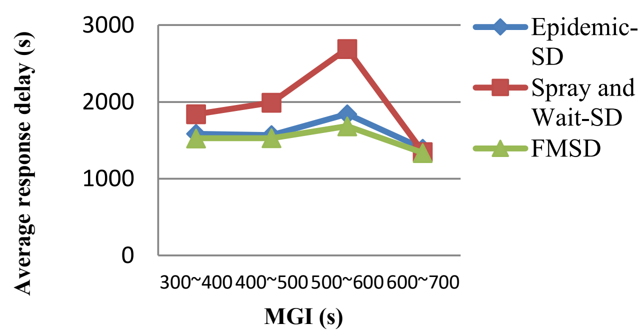

Figure 12. Average response delay for different MGI

As shown in Fig. 12, the average response delay of FMSD for different MGI is less than Spray and Wait-SD and Epidemic-SD.

In conclusion, comparing with Epidemic-SD and Spray and Wait-SD, the average recall and precision of FMSD are both improved certainly for different number of nodes or MGI without sacrificing too much time. And the average response delay of FMSD is less than the average request delay.

VII. CONCLUSION

This paper proposes a P2P service discovery algorithm based on Markov to improve both of the recall rate and precision in small delay, in order to adapt to the characteristics of high mobility, low density in the Internet of Things. Experimental analysis confirms the effectiveness of the proposed mechanism.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (No. 61640013), Natural Science Foundation of Inner Mongolia(No.2014BS0601,No. 2014BS0603),Program of Higher-level talents of Inner Mongolia University (No. 135138), the Inner Mongolia Autonomous Region Science and Technology Innovation Guide Reward Fund Project under Grant(No.20121317)and National Training Programs of Innovation and Entrepreneurship for Undergraduate(No.201510126034).The authors wish to thank the anonymous reviewers for their helpful comments in reviewing this paper.

REFERENCES

[1] A. Coman, M. A. Nascimento, et al. A framework for spatio-temporal query processing over wireless sensor networks. In Proceeedings of the 1st international workshop on Data management for sensor networks: in conjunction with VLDB (DMSN '04). Toronto, 2004, pp. 104-110.

[2] D. Chakraborty, A. Joshi, Y. Yesha, et al. GSD: A novel group-based service discovery protocol for MANETS. In Proceedings of the 4th International Workshop on Mobile and Wireless Communications Network, 2002, pp. 140-144.

[3] V. Navda, A. Kashyap, S. R. Das. Design and evaluation of imesh: an infrastructure-mode wireless mesh network. In Proceedings of the Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks(WoWMoM 2005). Giardini: Naxos, 2005, pp. 164-170.

[4] M. Nidd. Service discovery in DEAPspace. IEEE Personal Communications,8(4), 2001, pp. 39-45.

[5] G. Parisi. Field-theoretic approach to second-order phase transitions in two-and three-dimensional systems. Journal of statistical physics, 23(1), 1980, pp. 49-82.

[6] F. Sailhan, V. Issarny. Scalable service discovery for MANET. In Proceedings of Third IEEE International Conference on Pervasive Computing and Communications(PerCom 2005).2005, pp. 235-244.

[7] Z. Gao, Y. Yang, J. Zhao, et al. Service discovery protocols for MANETs: a survey. Mobile Ad-hoc and Sensor Networks, Springer Berlin Heidelberg, 2006, pp. 232-243.

[8] J. A. Garcia-Macias, D. A. Torres. Service discovery in mobile ad-hoc networks: better at the network layer?. In Proceedings of International Conference Workshops on of Parallel Processing(ICPP 2005 Workshops). 2005, pp. 452-457.

[9] M. Krebs, K. H. Krempels, M. Kucay. Service discovery in wireless mesh networks. In Proceedings of IEEE Wireless Communications and Networking Conference (WCNC 2008). Las Vegas, NV, 2008, pp. 3093-3098.

[10] S. A. H. Seno, R. Budiarto, T. C. Wan. Survey And New Approach In Service Discovery And Advertisement For Mobile Ad Hoc Networks. International Journal of Computer Science and Network Security (IJCSNS), 7(2), 2007, pp. 275-284.

[11] C. N. Ververidis, G. C. Polyzos. Routing layer support for service discovery in mobile ad hoc networks. In Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications Workshops(PerCom2005Workshops). Kauai Island: HI, 2005, pp.258-262.

[12] A. Vahdat, D. Becker. Epidemic routing for partially connected ad hoc networks. Technical Report CS-200006, Duke University, 2000.

[13] T. Spyropoulos, K. Psounis, C. S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In Proceedings of ACM SIGCOMM 2005 Workshops: Conference on Computer Communications. Philadelphia: PA, 2005, pp.252-259.

[14] A.Tiwari, N.Wadhawan, N.Kumar. A New Markov Chain Based Cost Evaluation Metric for Routing in MANETs. Communications in Computer & Information Science, 2010, pp.259-262.

[15] A. Sleman, R. Moeller. Integration of wireless sensor network services into other home and industrial networks; using device profile for web services(DPWS). In Proceedings of the 3rd International Conference on Information and Communication Technologies: From Theory to Applications(ICTTA 2008), Damascus, 2008, pp. 1-5.