2013 IEEE 27th International Conference on Advanced Information Networking and Applications

HBPR: History Based Prediction for Routing in Infrastructure-less Opportunistic Networks

Sanjay K. Dhurandher[1], Deepak Kumar Sharma[2], Isaac Woungang[3], and Shruti Bhati[1]

[1]CAITFS, Division of Information Technology, Netaji Subas Institute of Technology, University of Delhi, New Delhi, India.
Email: dhurandher@rediffmail.com, sb.shrutibhati@gmail.com
[2]Division of Computer Engineering, Netaji Subas Institute of Technology, University of Delhi, New Delhi, India.
Email: dk.sharma1982@yahoo.com
[3]Department of Computer Science, Ryerson University, Toronto, ON., Canada.
Email: iwoungan@scs.ryerson.ca

***Abstract*— In Opportunistic Networks (OppNets), the existence of an end-to-end connected path between the sender and the receiver is not possible. Thus routing in this type of networks is different from the traditional Mobile Ad hoc Networks (MANETs). MANETs assume the existence of a fixed route between the sender and the receiver before the start of the communication and till its completion. Routes are constructed dynamically as the source node or an intermediate node can choose any node as next hop from a group of neighbors assuming that it will take the message closer to the destination node or deliver to the destination itself. In this paper, we proposed a novel History Based Prediction Routing (HBPR) protocol for infrastructure-less OppNets which utilizes the behavioral information of the nodes to find the best next node for routing. The proposed protocol was compared with the Epidemic routing protocol. Through simulations it was found that the HBPR performs better in terms of number of messages delivered and the overhead ratio than the Epidemic protocol.**

Keywords: Opportunistic network (Oppnets), Opportunistic routing, The ONE (Opportunistic Network) simulator.

I. INTRODUCTION

Opportunistic Networks [1] are the variants of *Delay Tolerant Networks* (DTNs) [2, 3]. It is a class of networks that has emerged as an active research subject in the recent times. Owing to the transient and un-connected nature of the nodes, routing becomes a challenging task in these networks. Sparse connectivity, no infrastructure and limited resources further complicate the situation. Hence, the challenges that one is likely to face while routing in opportunistic networks is very different from the traditional wireless networks. However, their utility and potential for scalability makes them a huge success. These networks can be useful for routing in places where one is not likely to find base stations and connected routes for long distances. Examples include extreme terrains, mountains, grasslands and space communication.

OppNets use hop by hop methodology to deliver message to its destination. Unlike MANETs, OppNets do not have to worry about finding or creating an end to end connected path. Hence, the protocols meant for MANETs cannot be used here. OppNets can deliver messages even though there might be no knowledge of a previous path [4]. They study and utilize other aspects, behavior and characteristics of nodes to deliver the messages. High mobility and frequent disconnections are considered to be norms rather than hindrances. Devices are often out of range of a network but may be connected to other devices. They utilize these meetings and connections for data delivery.

The routing protocols used by the OppNets can be divided into two major categories namely Infrastructure-less protocols and Infrastructure-based protocols [5]. Infrastructure-less protocols make no previous assumptions regarding the nodes and the network topology. Every single node in the network behaves as a peer and there is no master-slave relationship among the nodes. No infrastructure existence is assumed which can help in the forwarding of the messages. These protocols are meant for the flat ad hoc networks.

The nodes in infrastructure-less OppNets are usually mobile devices carried by humans or vehicles. Hence they usually exhibit a pattern in which they move during a period of time. They have communities they might belong to or places where they visit with more frequency. Understanding the nodes' movements, communities and members of those communities can help us model a structure in the network. This previous knowledge of their movement can be used for better delivery of the messages.

Routing of packets in accordance with the community patterns is sure to give better results as according to sociology '*correlated interaction*' implies that organisms are more likely to interact with organisms of their own type rather than others [6]. In this paper, a novel approach the HBPR protocol is introduced to address the message delivery problem in infrastructure-less OppNets. It assumes Custom Human Mobility Model, which is discussed in Section IV, for the movement of the nodes in the network.

The rest of the paper is organized as follows. Section II presents the background and related work in the area of routing protocols in infrastructure-less OppNets. Section III describes the proposed protocol in detail. In Section IV, we discuss the simulation setup and various assumptions used to simulate the HBPR protocol. Section V is devoted to the simulation results, where we discuss the performance of the HBPR and compare it with the Epidemic routing protocol [7]. Finally, Section VI concludes our work and provides some insights on the future work.

1550-445X/13 $26.00 © 2013 IEEE
DOI 10.1109/AINA.2013.105

II. BACKGROUND AND RELATED WORK

Geographic Location information has been widely used earlier to implement effective routing in networks [8, 9]. They work very well in wireless ad hoc networks because they make networks easily scalable. The work in [10] is believed to be the oldest that use the Geographical Locations for this purpose. According to [11], Geographic routing is the practice where all the nodes know their own location and every node forwards a packet to its neighbor that is geographically closest to the destination, so long as that neighbor is closer to the destination. Typical geographical routing algorithms use planar graphs for this purpose. However, HBPR makes use of the history of node's movement to determine its proximity to the destination. To implement Geographical routing, a node must be aware of two things [12] – (a) its own location as well as its neighbor's location and (b) the source of the message must know the location of the destination.

Several routing protocols have been proposed in the past to address the problem of routing in OppNets such as Epidemic [7], PRoPHET [13], HiBOp [14] and CAR [15]. In this section, a brief overview of these routing protocols is given.

A. Epidemic

The Epidemic routing [7] protocol makes minimum assumptions about the topology and connectivity of the underlying network. It is a purely dissemination based routing protocol. Epidemic routing relies upon the networks capability to eventually deliver the message by opportune visits. Every node maintains a list of the messages that are contained in its buffer. This list is maintained in the form of a Summary Vector. Every node in the network has an identifier associated with it. When two nodes meet, the node with smaller identifier initiates a session called anti-entropy session. Recently met nodes are ignored by cross-referencing with a cache of recent meet-ups. Both the nodes exchange summary vectors and exchange the messages that the other doesn't have. As expected, this protocol suffers heavily from the overhead of extra copies of messages in the network. This might lead to dropping of messages from the message queue.

B. PRoPHET

PRoPHET (Probabilistic Routing Protocol using History of Encounters and Transitivity) [13] uses the history of encounters and transitivity to deliver the message assuming that nodes move in a predictable fashion and not randomly. The delivery predictability P(a,b) is the probability with which node A is likely to meet node B in the future. If the neighbor has more probability of meeting the destination node, the carrier node transfers the message to the neighbor.

C. HiBOp

This HiBOp (History Based Routing Protocol for Opportunistic Networks) [14] utilizes a node's present context to find a better path for the faster delivery of the message. Current context of a node can be thought of as a snapshot of the environment it currently resides in. Two tables are used – *Identity Table (IT)* and *History Table*. It determines whether there is a match between the context information of the node and the information associated with the message and then forwards the message to that node.

D. CAR

CAR (Context Aware Routing) [15] relies on the nodes' '*logical connectivity information'* to decide the next hop for the message. The protocol first decides if the message recipient is in the same cloud or not. If yes, it uses *DSDV* [16] protocol to deliver the messages. Other relay nodes are chosen such that they have a higher probability of message delivery which is calculated at every node. Apart from this the protocol also predicts the future values of the context so as to make the delivery more effective.

III. THE HBPR PROTOCOL

In this section, the HBPR protocol is introduced in detail. The following assumptions are taken into account while designing the protocol: (1) The nodes in the network move according to the Human Mobility pattern; (2) The nodes are cooperative and do not operate with malicious instincts; and 3) The area in which the protocol is simulated is divided into cells. Each cell is given a number. These cells can be locations that nodes visit in the course of the simulation or their home cells. While storing the tables, we store the cell numbers instead of exact coordinates. The protocol is divided into three phases – (a) Initializing the Home Locations, (b) Message generation and Home Location update and (c) Next hop selection.

A. Initializing the Home Locations

It is assumed that nodes behave as in a human mobility model. Each node has a certain location that it might visit more frequently, some that it might visit rarely and so on. Based on this, one can predict a node's future location quite easily using its location's past history. Since in geographical routing it is required that the destination's GPS location be known to us, it is only plausible that the destination flood that information in the network. However such a scenario would result in network clogging. Therefore, keeping in mind that the network follows Human Mobility model, nodes are allowed to flood the network with their most visited location when the network becomes operational.

A head start is given to the network before messages are spread in order for their history to get settled and a pattern to emerge. During the course of operation, if a node changes its pattern and a different location has more frequency, it floods the network with this new information and a time stamp to distinguish the new *Home Location* information from its old information. Thus, when a source to destination (SD) line is drawn, it joins the source/ current node carrying the message with the destination's most visited location. In this way we target a node's behavior instead of its current state making our algorithm more suitable for a Human mobility scenario.

B. Message Generation and Home Location update

This phase has two parts. The first is the generation of new messages at some of the nodes. The destination node's ID is obtained from the message itself. The second part is concerned with updating the *Home Location Table* which is described later. This part helps us to adjust according to the changing relationship and behavior of the nodes. However, to account for the changing movements of the node, every node must track its own history. After a fixed time interval (*refresh_period)* nodes re-scan history to figure out their *Home Location* cell. It should be noted that the *refresh_period* must be large enough so as not to incur calculation overheads.

C. Next hop selection

A *Utility Metric* is used in order to decide the next hop for the message, which is obtained as a function of three parameters –

1) Stability of node's movements: In order to predict the stability of node's movement *S*, the *History Table* is used. As the node moves, it records its coordinates as well as the time. Thus, a node can utilize this information to calculate its own average speed over two different positions. We have a list of such average speeds. Using this list, it can be analyzed whether the change in average speeds is very large or nominal. A large change signifies an unstable movement whereas a nominal change accounts for a stable node. Initially all the nodes are given a stability value equal to zero. For every two consecutive speeds if the change is greater than 10 units per second, the stability is decreased using the formulae:

$$S = S - (1-S)*S_{int} \qquad (1)$$

It remains constant if the change is less than 10 units per second.

2) Prediction of the direction of future movement using Markov Predictors: Markov models can be used to predict the next location based on the histories. These models use the past few locations to predict the next one. The length of the past locations is known as the order of the *Markov Predictor* [17, 18]. A *Markov Predictor* is a simple *Markov Model* that can be used for this purpose. It maintains a table with the frequencies of visit for every location for the given pattern of visits. This is then used to predict the next location. For example, if the past history is
AGHBGTYGHIGHYKLOPWNGHWBKJDNGHRJBFJ
GHYKJFNGHYLKJNKSGHWOKSADGH

Then, according to a Markov predictor of order 2, the next location would be Y because Y occurs the most number of times as is shown below

A**GH**BGTY**GH**I**GH**YKLOPWN**GH**WBKJDN**GH**RJBFJ
GHYKJFN**GH**YLKJNKS**GH**WOKSAD**GH**

3) The perpendicular distance of the neighboring nodes from the line of sight of source and destination (SD line): This metric helps us in selecting nodes that are on an average at a closer distance to the SD line because they have to travel a lesser distance than the nodes that are away from it.

Using these three aforementioned parameters the *Utility Metric* denoted by *U(i)* of the i^{th} node is calculated with the formula:

$$U(i) = \sum_{j=1}^{j=3} W(j)*V_i(j) \qquad (2)$$

where *W(j)* is the weight of the j^{th} parameter and $V_i(j)$ is the value of the j^{th} parameter for i^{th} node i.e. *Vi(1)* is the Stability metric, *Vi(2)* is Prediction metric and *Vi(3)* is the Perpendicular distance metric for node *i*. *U(i)*is thus calculated for node *i* and a threshold *T* can be used to decide its selection as the next hop for the message. The message is then forwarded to those neighboring nodes that have a *U(i)* value greater than *T*. Thus, *T* is used to control the amount of flooding in the network. The HBPR protocol uses two tables in the course of *Utility Metric* calculation. These tables are detailed below.

History Table

In order to predict next node location and calculate stability of a node, a record of the coordinates (location) and timestamp is required. At the start of the simulation, every node has its timer set to zero and at time interval t_{in} the location (coordinates in this case) and time stamp is stored in the internal *History Table* of every node. At any point in time, the node maintains a record of 100 locations. The 101^{th} location is automatically deleted by the HBPR to account for the changing behavior in the node's movement and to conserve the memory space. Table1 shows a snapshot of the history table at a particular instant of time.

Table1: History Table

Time	Location
0.45	23
1.2	45
2.6	45
2.9	40
3.3	45

Home Location Table

Every node maintains a table where it stores the *Home Location* for other nodes in the network. A sample snapshot of the *Home Location Table* at a particular instant of time is shown in Table2. At every encounter, the nodes exchange and update these tables. For same node entries, the time field is matched to determine whether or not to update the entry. Only the most recent Home Location entry is considered and others are discarded. This can help nodes to keep up with their changing movement pattern.

Table2: Home Location table.

Host ID	Home Location
A1	68
B3	43
H7	45
F3	12
A5	6

Figure1 shows the example of a network scenario used to simulate the working of HBPR protocol.

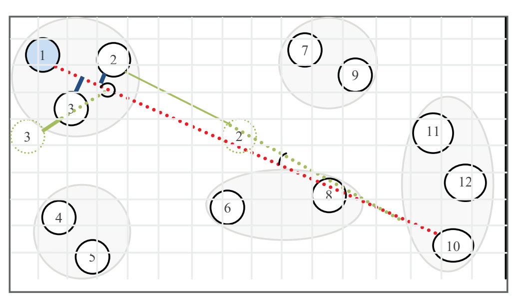

Figure1: An example scenario of the HBPR protocol.

In Figure1, the red line represents the SD (source to destination line), blue line represents the perpendicular distance from the SD line, and the dotted green circles represent the future predicted locations of the nodes. The source is node 1 and destination is node 10. They are connected by an imaginary SD line. Node 2 and node 3 are in the range of node 1 and their future predicted locations are shown using dotted green circles. The future and present locations is connected using a straight line and its angle with the SD line is calculated. The perpendicular distance with the SD line is shown using dark blue line.

IV. Simulation Setup

The performance of HBPR is evaluated using the ONE simulator [19]. The nodes are mobile and have been divided into six groups and each has 25 nodes. The first and third group nodes are of pedestrians with speed between 0.5 – 1.5 m/sec. The second group is of cyclists with speed varying between 2.7 – 13.9 m/sec. The fourth, fifth and sixth group nodes are cars with speeds varying from 7 – 10 m/s. The first, second and third group have the same *Home Location* while the fourth, fifth and sixth groups have different *Home Locations.* The mobile nodes have a transmission range of 10 meters and transmit speed of 2 Mbps. Each simulation is run for 43000 seconds. The world size for the movement model is 4500m x 3400m meters. A new message is generated at every 25 – 35 seconds and the size varies from 500 KB to 1 MB. For the results in the next section, the following values of weights i.e. *W(j)* are used for the three different parameters. The Stability Metric weight *W(1)* = 0.4, the Prediction Metric weight *W(2)* = 0.4, and the Perpendicular distance metric weight *W(3)* = 0.2.

The whole world size is divided into cells of 100m x 100m. A new *Custom Human Mobility Model* for the HBPR protocol has been implemented which is inspired from [13]. The nodes are then grouped into six communities. Each community has a *Home Location* cell. At the start of the simulation, every node is present in its *Home Location* which it disseminates through the network using its own *Home Location* table. The node travels to its *Home Location* with a probability *p* and to all other locations with probability *1-p*. The use of this mobility model can be attributed to the fact that HBPR is designed to perform best with Human scenarios. It works on the existence of community like structures and a recurring pattern. Hence, its performance will surely decrease in a movement model where the nodes move randomly rather than with a predetermined destination.

The following settings and configurations have been used in the simulation:

1) *Varying the threshold:* The threshold is varied to observe the performance of the HBPR protocol.
2) *Varying the number of nodes in the simulation:* The total no of nodes in the simulation are taken as 150, 180, 210, 240 and 300 to compare HBPR against the Epidemic routing protocol.

The performance metrics used are:

a) *Total no of messages delivered:* It is total number of messages received by the destination.
b) *Overhead ratio.* It is the average number of forwarded copies per message.

V. Simulation Results

A. Varying the threshold

While varying the threshold *T,* the number of nodes is kept fixed to 240 and the *T* is varied from 0.2 to 0.6 with an increment of 0.1 each time. If a node has *Utility Metric* value greater than *T* the message is transferred to it.

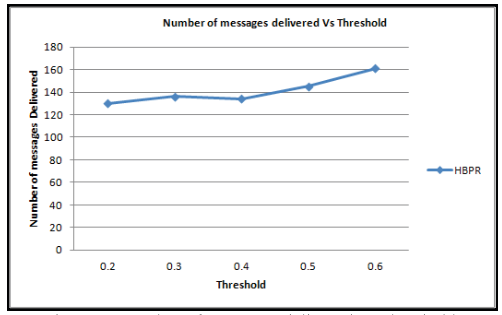

Figure 2: Number of messages delivered vs. threshold.

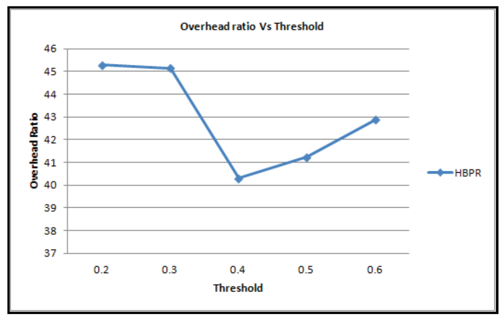

Figure 3: Overhead ratio vs. threshold.

In Figure 2, the number of messages delivered gradually increase as the threshold is increased. This is due to the fact that only better relay nodes are chosen for delivering the message (i.e. nodes that have a higher *Utility Metric*).This ensures that flooding is controlled and lesser messages get dropped thus showing an increase in the number of delivered messages. In Figure 3, the overhead ratio decreases initially when the threshold is increased but then it starts increasing. It is observed from the graph that a threshold of 0.6 is more suitable for this set of simulations as it gives a high number messages delivered while keeping the overhead ratio low.

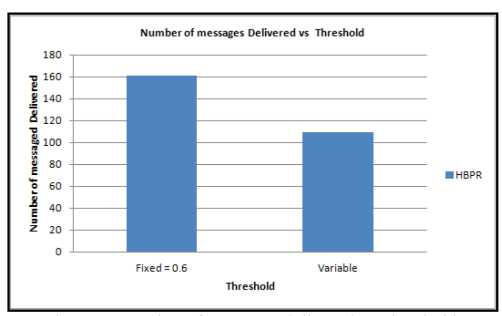

Figure 4: Number of messages delivered vs. threshold.

Figures 4 and 5 represent the cases when the threshold is taken as *'fixed'* and *'variable'* along the X- axis. In case of *'fixed'*, the threshold is taken to be 0.6 and the message is transferred from the current node to all those neighbor nodes which have *Utility Metric* greater than 0.6. In case of *'variable'* threshold, a node *i* currently carrying the message will forward it to the neighbor node *j* only if the *Utility Metric U(j)* of node *j* is greater than the *Utility Metric U(i)* of node *i.* The performance of the *variable* threshold is compared with the *fixed* threshold next.

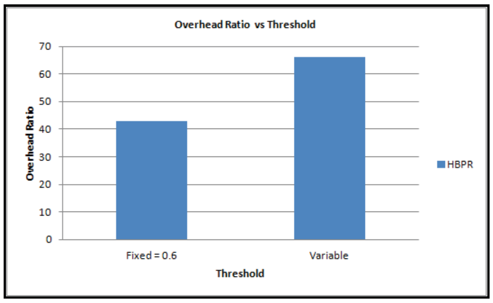

Figure 5: Overhead ratio vs. threshold

It can be seen in the graphs of Figure 4 and 5 that the performance of the variable threshold suffers heavily both in terms of number of messages delivered and overhead ratio. This is due to the fact that for situations where the carrier node *i* has a lower value of *U(i)* than most of its neighbors, it still transfers messages even if none of the nodes are good relay nodes. This leads to a flooding-like situation where the overhead ratio increases and also leads to dropping of messages, thus a lesser number of messages are delivered.

B. Comparison of HBPR with Epidemic routing Protocol

Figures 6 and 7 provide a comparison of the performance of the *HBPR* protocol against the *Epidemic* routing protocol using the *Custom Human Mobility Model*. The simulation is performed with a fixed threshold value of 0.6.

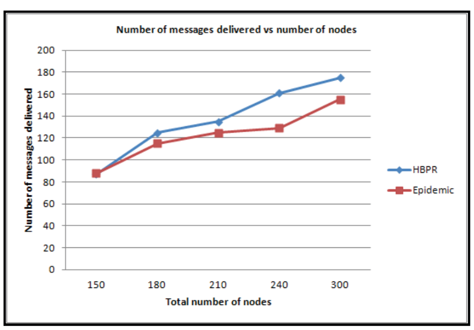

Figure 6: Number of messages delivered vs. number of nodes.

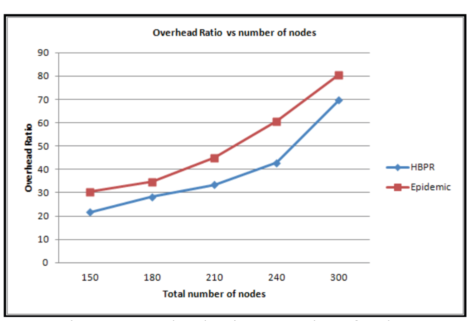

Figure 7: Overhead ratio vs. number of nodes.

Figures 6 and 7 show that the HBPR protocol has a significantly lower overhead ratio and more number of messages delivered as compared to the Epidemic routing protocol. Thus, it helps the nodes to achieve better bandwidth. This can be attributed to the fact that the node makes intelligent decisions through the calculation of the *Utility Metric* before relaying the message. Whereas, in the case of Epidemic routing, copies of messages are blindly flooded. This results in greater overhead as well as dropping of several copies. Therefore, the number of messages delivered in case of Epidemic routing also falls.

VI. CONCLUSION AND FUTURE WORK

In this work, we proposed an algorithm for routing in infrastructure-less Opportunistic Networks. The protocol has been inspired from the Geographical Routing technique and uses history of movement to model the behavior of the nodes. Using Markov Predictor, prediction is carried out to select the best next hop node. The proposed protocol was observed to perform well in terms of the number of messages delivered and the overhead ratio for different threshold values. Further, in terms of these performance metrics, the proposed HBPR protocol was also found to outperform the Epidemic routing protocol.

In future, we plan to introduce message acknowledgements in the HBPR protocol and compare it with some other existing routing techniques in infrastructure-less Opportunistic Networks. We will also evaluate the performance of the HBPR protocol with help of some other metrics like delivery probability, average delay etc. by varying other parameters like TTL and the speed of the nodes.

REFERENCES

[1] L. Lilien, Z.H. Kamal, V. Bhuse, and A. Gupta, "Opportunistic Networks: The Concept and Research Challenges in Privacy and Security," *in Proceedings Of NSF Intl. Workshop on Research Challenges in Security and Privacy for Mobile and Wireless Networks (WSPWN 2006)*, Miami, March 2006, pp. 134-147.

[2] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets", *in proceedings of ACM SIGCOMM 2003*, Karlsruhe, Germany, 25-29 August, 2003, pp. 27-34.

[3] S. Jain, K. Fall, and R. Patra, "Routing in a Delay Tolerant Network", *in proceedings of ACM SIGCOMM 2004*, 30 Aug.-3 Sept. 2004, pp. 145–158.

[4] S. K. Dhurandher, D. K. Sharma, I. Woungang, and H.C. Chao, "Performance Evaluation of Various Routing Protocols in Opportunistic Networks", *in Proceedings of IEEE GLOBECOM Workshop 2011*, Houston, Texas, USA , 5-9 December, 2011, pp. 1067-1071.

[5] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks", *IEEE Communications Magazine*, vol. 44, Issue 11, November 2006, pp. 134-141.

[6] S. Okasha. Altruism, group selection and correlated interaction. *British Journal for the Philosophy of Science*, 56(4):703–725, December 2005.

[7] A. Vahdat, and D. Becker, "Epidemic routing for partially connected ad hoc networks", *Technical Report CS-2000-06, Dept. of Computer Science, Duke University*, Durham, NC, 2000.

[8] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks" *in proc. of* the 3rd international workshop on Discrete algorithms and methods for mobile computing and communications *(DIALM '99)*, Seattle, WA, USA, 20 Aug. 1999, pp. 48–55.

[9] J. Gao, L. Guibas, J. Hershberger, L. Zhang, and A. Zhu, "Geometric spanner for routing in mobile networks*", in proc. of the 2nd ACM international symposium on Mobile ad hoc*

networking & computing (MobiHoc '01), Long Beach, CA, USA, 04-05 October 2001, pp. 45–55.

[10] H. Takagi, and L. Kleinrock, "Optimal transmission ranges for randomly distributed packet radio terminals". *IEEE Transactions on Communications,* Vol. 32, Issue 3, March 1984, pp. 246–257.

[11] Y.J. Kim, R. Govindan, B. Karp and S. Shenker, "Geographic Routing Made Practical", in p*roceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation* - Vol 2, CA ,USA, 2005, pp. 217 – 230.

[12] F. Kuhn,R. Wattenhofer and A. Zollinger, "Asymptotically optimal geometric mobile ad-hoc routing", *in proceedings of the 6th international workshop on Discrete algorithms and methods for mobile computing and communications (DIALM '02),* Atlanta, GA, USA, 28 September 2002, pp. 24 – 33.

[13] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks", *ACM SIGMOBILE, Mobile Computing and Communications Review*, vol. 7, Issue 3, July 2003, pp. 19–20.

[14] C. Boldrini, M. Conti, I. Iacopini and A. Passarella, "HiBOp: a History Based Routing Protocol for Opportunistic Networks", *in proceedings Of IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks, 2007 (WoWMoM 2007)*, Espoo, Finland, 18-21 June 2007, pp. 1-12.

[15] M. Musolesi,S. Hailes and C. Mascolo. "Adaptive Routing for Intermittently Connected Mobile Ad Hoc Network". in proceeding of Sixth IEEE International Symposium on World of Wireless Mobile and Multimedia Network (WOWMOM '05), Taormina-Giardini Naxos, Italy, 13-16 June 2005, pp 183-189.

[16] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computer*", in proceedings of the conference on Communications architectures, protocols and application (SIGCOMM'94),* London UK ,August 1994, pp. 234-244.

[17] I.-C. K. Chen, J. T. Coffey, and T. N. Mudge, "Analysis of Branch Prediction via Data Compression", *in proceedings of the seventh international conference on Architectural support for programming languages and operating systems (ASPLOS VII)*, Cambridge, Massachusetts, USA, 1-4 October 1996, pp. 128–137.

[18] S. M. Ross. Introduction to Probability Models. *Academic Press,* 1985.

[19] A. Keranen. " Opportunistic Network Environment Simulator", *Special Assignment Report, Helsinki University of Technology, Dept. of Communications and Networking*, May 2008.