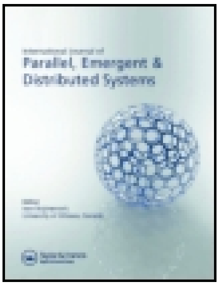

International Journal of Parallel, Emergent and Distributed Systems
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gpaa20

A direction-based geographic routing scheme for intermittently connected mobile networks

Ze Li [a] & Haiying Shen [a]

[a] Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, 29631, USA

Published online: 08 Jan 2013.

International Journal of Parallel, Emergent and Distributed Systems
iFirst article, 2012, 1–26

Taylor & Francis
Taylor & Francis Group

A direction-based geographic routing scheme for intermittently connected mobile networks

Ze Li[1] and Haiying Shen*

Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29631, USA

(*Received 13 September 2012; final version received 13 September 2012*)

In an intermittently connected mobile network, a complete routing path from a source to a destination cannot be guaranteed most of the time. Therefore, traditional routing methods for mobile ad hoc networks are not applicable in such a network. Current approaches for intermittently connected mobile networks are primarily based on redundant transmission and single-copy opportunistic routing. However, they incur either high overhead due to excessive transmissions, or long delay due to incorrect path choices during forwarding. In this paper, we propose a direction-based geographic (DIG) routing scheme for intermittently connected mobile networks. Relying on geographic location information, the packets are routed in a path approximately to the shortest path from the source node to the destination, which significantly reduces the overhead in redundant transmission and decreases the transmission delay in the single-copy opportunistic routing. Theoretical analysis and trace-driven experimental results show that DIG provides low transmission delay with low overhead in comparison with the schemes in the redundant transmission and single-copy opportunistic routing.

Keywords: delay tolerant networks; single-copy routing; direction-based geographic routing

1. Introduction

With the rapid development of wireless communication and electronic techniques, mobile devices are quickly growing in their communication capabilities, which makes the concept of ubiquitous computing very promising in the near future. One research area that currently receives increasing attention is mobile ad hoc networks (MANETs). MANETs are collections of wireless mobile nodes which promise a convenient infrastructure-free communication. In the absence of a central control infrastructure, the hosts in a MANET communicate with each other in a multi-hop fashion [15,33], in which a continuous connection between a source node and a destination node should be guaranteed.

One special group of MANETs is Delay Tolerate Networks (DTNs) (i.e. intermittently connected mobile networks) in which source nodes and destination nodes are intermittently connected. Examples of DTNs include wildlife monitoring sensor networks [16], interplanetary communication networks [4], vehicular ad hoc networks [46], terrestrial wireless networks and ocean sensor networks [28,32]. The intermittent connectivity in these networks is caused by node mobility [46], power management [16], wireless transmission range, sparsity [14] or malicious attacks [6]. Therefore, conventional Internet routing protocols (e.g. RIP [38] and OSPF [30]) as well as MANET routing schemes, such as DSR

*Corresponding author. Email: shenh@clemson.edu

ISSN 1744-5760 print/ISSN 1744-5779 online

http://dx.doi.org/10.1080/17445760.2012.731057
http://www.tandfonline.com

[15] and AODV [33], that try to discover continuously connected paths with the minimum transmission cost before sending data out are not applicable here.

Currently, almost all routing schemes proposed for DTNs are topology-based routing, in which the routing is conducted based on node connectivity in the network. The schemes can be classified into two categories: flooding-based routing (e.g. epidemic routing) [16,27,40] and single-copy opportunistic routing [39]. Despite of their high robustness and low transmission delay, flooding-based routing schemes consume much energy, bandwidth and memory space that are crucial to wireless network applications. In the single-copy opportunistic routing such as two-hop direct routing, a source node transmits each segment of the source packet stream to a relay node. The packets are allowed to be buffered for a long time until the relay nodes meet the destination node. Although such a scheme brings about much lower overhead for packet transmission, it suffers from severe transmission delay if a node chooses a wrong path for the delivery.

Geographic routing has been proposed for MANETs which relies on geographic position information of mobile nodes instead of using the node connectivity information. Each node in geographic routing only maintains the location information about its current neighbour nodes, which is generated by Global Position System (GPS) or numerable virtual coordination methods [8,34]. One requirement of geographic routing is that the source node should be aware of the location of the destination node. Fortunately, in the majority of applications of DTNs such as wildlife monitoring sensor networks [16] and interplanetary communication networks [4], the positions of the destination nodes (sinks) are determined and can be easily known by all the nodes in the system. It can also be realised by periodically broadcasting the location information of the destination. The packets are routed to these location-determined sinks for data collection, data processing or further transmission to clients via Internet. Although geographic routing methods generate much less transmission overhead and have high transmission scalability for decentralise routing, current geographic routing methods proposed for the MANETs [3,18] using a greedy transmission strategy are not applicable to DTNs. In MANETs, the packets are greedily transmitted to the destination via the continuously connected link based on geographic locations of intermediate neighbour nodes. However, because of the long delay between two transmission steps in DTNs, a node currently close to the destination node cannot be guaranteed to be close to the destination later or can forward the packet to a node closer to the destination in the near future. That is, no node can expect when and which nodes it will meet in the near future. Therefore, the traditional geographical routing methods for MANETs, which only emphasise current locations of neighbour nodes, are not applicable in DTNs.

In this paper, we present the design, implementation and evaluation of a direction-based geographic (DIG) routing scheme for DTNs. DIG aims to quickly forward packets to the destination nodes. It tries to forward packets to the relay nodes that have long communication time with the destination node, thus increasing the number of packets that can be transmitted to the destination in one interaction. Specifically, in DIG, the source node initially forwards its packets to the neighbour node that locates closer to the destination node regardless of its moving direction. Then, if the relay node's distance to the destination is longer than a pre-defined threshold, the node greedily forwards packets to nodes that move towards and are currently located closer to the destination, in order to forward the packets towards the destination quickly. Otherwise, the node forwards packets to the nodes that have a longer predicted communication time regardless of their locations, in order to ensure a long communication time between a relay node and the destination node. Once the relay node moves into the transmission range of the destination node, it continuously transmits the packets to the destination node until the connection breaks.

Buffer refers to the memory in which a mobile node stores received packets. DIG also has a buffer management strategy to reduce average packet transmission delay. We build a module to theoretically analyse the performance of DIG. Theoretical and simulation results show that:

(1) With a large buffer size, DIG generates delay comparable to the epidemic routing scheme, but with much lower communication overhead.
(2) With a small buffer size, DIG yields significantly shorter delay, lower overhead and higher throughput than the epidemic routing scheme.
(3) Compared to the single-copy opportunistic routing scheme, DIG produces much lower delay with comparatively high throughput.

In the next section, we discuss existing related work. Section 3 presents the DIG routing scheme. In Section 4, we analyse the performance of DIG theoretically. Simulation results are presented in Section 5. Finally, Section 6 concludes the paper.

2. Related work

Although numerous routing protocols for MANETs have been proposed [5,15,33], they are not appropriate for DTNs that are sparse and intermittently connected [39]. An initial method to deal with connectivity disruptions in DTNs is to reinforce connectivity on demand by filling the 'communication gap' with specialised nodes (e.g. robots and satellites) when a disconnection occurs [25,48]. However, such a method needs global monitoring of the network and is not applicable to self-organising DTNs.

One category of proposed methods is flooding-based routing [37,43–45,47]. Epidemic routing [43] is the most intuitive flooding routing scheme in DTNs, in which packets are forwarded between two nodes when they come into contact with each other until the packets arrive at their destinations. This scheme can guarantee a short delay by locating a shortest routing path at the cost of high network resource consumption. There are a number of improved approaches proposed to reduce the overhead of the epidemic routing [37,44,45,47]. Zhang et al. [47] proposed to let packets be forwarded to partial neighbours instead of flooding in order to reduce the transmission overhead in the epidemic routing. Small et al. [37] proposed a buffer management scheme which tries to remove redundant copies of packets in packet forwarding. Wang et al. [44] and Widmer et al. [45] proposed to use the network-coding scheme to reduce the overhead of the epidemic routing. The source packets are coded and a fraction of the generated code-blocks are distributed to the nodes. This reduces the overhead of transmitting a full copy of the source packets. Although these schemes can improve the performance of the epidemic routing to a certain extent, they still inherit the high overhead shortcoming of flooding and cannot significantly reduce transmission delay.

Another category of routing schemes for DTNs is single-copy opportunistic routing [9,16,27,12]. In [16], nodes record the history of its past encounters in order to make fewer but more informed decisions. Those routing paths are predicted either by statistics of a mobility module or by a historical moving path record. However, these schemes reduce the transmission overhead of flood-based routing at a significant penalty on delivery delay. Ferriere et al. [9] proposed to consider encounter ages to improve the route discovery process of regular ad hoc networks. Instead of searching for the destination, a source node searches for any intermediate node that has higher encountering frequency with the destination node. Lindgren et al. [27] pointed out that consulting the age of the last node encountered when making forwarding decision results in superior performance than flooding. Grossglauser and Tse [12] proposed a direct transmission scheme to increase the capacity of the MANETs. In the scheme, a source node forwards its packets to a certain number of relay nodes, which hold these packets until meeting the destinations.

Many geographic routing schemes [2,10,11,13,17,20,22,23,29,41] have been proposed in MANETs. In the geographic routing, relying on GPS, a node knows the geographic locations of the destination and its neighbours and forwards a message to the node geographically closest to the destination. Hou and Li [13] and Takagi and Kleinrock [41] perhaps proposed the first geographic routing schemes. In the schemes, a node greedily routes a message to its neighbour node that is the closest to the destination. However, the geographic routing may fail if a message reaches a node which does not have a neighbour geographically closer to the destination. To deal with this problem for message delivery guarantee, Kranakis et al. [20] proposed a face routing, in which a message walks along faces of planar graphs and proceeds along the line connecting the source and the destination. As the faces of planar are continuous from source nodes to destination nodes, the messages will not halt at nodes just because there is no closer neighbour to the destination node. Later on, several other geographic routing mechanisms were proposed for guaranteed message delivery [2,11]. In the mechanisms, the entire network is partitioned into faces that are bounded by polygons made up of edges of the network. The messages are routed along the node in the boundaries of the faces. However, these routing protocols fail to provide better worst case performance than the original face routing. Kuhn et al. [22] proposed a Greedy Other Adaptive Face Routing (GOAFR) mechanism. Instead of changing to the next face at the 'best' intersection of the face boundary between the source node and the destination node, GOAFR changes to the boundary point that closest to the destination node. However, the face routing and GOAFR are not applicable for practical purposes since it is costly to partition a network into plantar faces. Therefore, Karp et al. [17,19,21] proposed to combine face routing with greedy routing, in which messages are initially forwarded greedily. When the messages cannot be forwarded further, the face routing is employed. Leong et al. [23] and Frey and Stojmenovic [10] further proposed to improve the geographic routing by letting nodes store some aggregated information of nodes in the local areas to assist the geographic routing. Such local area information can provide routing information to decide which direction in the routing paths is most likely to make progress towards a given geographic destination. In general, all these routing mechanisms can only apply to the scenario in which nodes are fully connected, but are not applicable in DTNs.

3. DIG routing scheme

Like the traditional geographic routing, we assume each node knows its current location and moving direction using GPS or virtual coordinators [8,34]. Traditional geographical routing cannot be directly used in DTNs because current closest nodes to the destination are not necessarily the closest node later on. DIG aims to forward packets to nodes that will meet the destination earlier and transmit more packets to the destination when meeting the destination. We present the desirable design goals of DIG and corresponding strategies it uses below.

(1) In order to reduce transmission overhead and avoid packet congestion, rather than relying on flooding, DIG uses a single copy of the packets in routing.
(2) In order to reduce transmission delay, unlike existing single-copy opportunistic routing schemes [35,36], DIG uses the node location and mobility information to forward packets to nodes that have high probability to meet the destinations and long transmission time with the destinations.

(3) In order to increase delivery throughput and reduce average transmission delay, DIG has an efficient buffer management strategy.

In this paper, we adopt the Uniform Mobility Model (UMM) [1] as our theoretical analysis scenario. In this model, each of the n mobile nodes moves at speed v inside a unit of a circular disk. At time $t_0 = 0$, the positions of these nodes are distributed uniformly at random inside the disk. Also, the directions of the motion of the n nodes are identical and independently distributed (i.i.d.) in $[0, 2\pi)$. After $t_0 = 0$, a node behaves as follows: it picks a direction uniformly at random from $[0, 2\pi)$ and moves in that direction for a certain distance at speed v.

We assume that the destination nodes are static or their locations are known to all the nodes in the system. Such assumption is realistic since most of the current applications [4,16,46] have location-determined data sinks. Also, such an assumption holds true in numerous real environment scenarios such as interplanetary networks, terrestrial networks and wireless senor networks. DTN nodes use certain storages such as disk and flash memory for storing packets. We use buffer to represent all kinds of storages.

3.1 Overview of the DIG scheme

In DTNs, a routing utility of a node, such as contact frequency and contact duration, is a measure of the node to increase the value of a routing metric such as throughput and delay [24]. DIG defines two routing utilities for the message routing: *relative angle* and *predicted communication time* between nodes and destination nodes, to predict the communication time between a node and a destination.

The basic idea of the algorithm is to route packets towards the destination quickly and try to guarantee that the packets are forwarded to the relay nodes that have a long communication time with the destination node. When a node is far away from the destination node, it is difficult to predict its communication time with the destination node, since the node may constantly change its moving direction. Therefore, in this case, we let the packet holding node forwards packets to the nodes that move towards and are currently located closer to the destination regardless of their predicted communication time to the destination. When a node is located close to the destination node, its moving direction relative to the destination does not greatly change because of its short distance to the destination node. Therefore, in this case, we let a node forward packets to the nodes having a long predicted communication time with the destination node. We set a threshold T for a node's distance to the destination to identify the two cases. After the node moves into the transmission range of the destination node, it will continuously transmit the packets to the destination node until the communication link breaks.

In DIG, two nodes exchange their current locations and moving directions when they meet each other. Each node then calculates the routing utility defined by DIG of the other node, and decides whether to forward the packets to the other node. Figure 1 shows an example of packet routing from a source node to a destination node in DIG. As the figure shows, when the packets are in relay node R_1, which is far from the destination node D (i.e. beyond a threshold distance T), the packets are greedily forwarded to the next hop node R_2 that is located closer to the destination and moves towards the destination. In this way, the packets can be forwarded towards D quickly.

When the packets are forwarded to the nodes that are located within a threshold distance T to the destination but not within the transmission range of the destination node, they are only forwarded to a mobile node R_3 which will have a long communication time

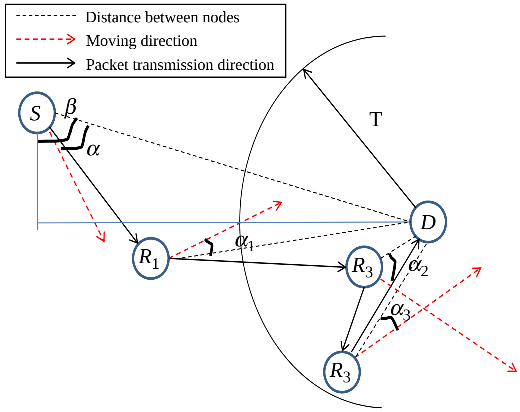

Figure 1. Transmission process.

with the destination node, even if the distance between R_3 and D is long. In this way, we can guarantee that more packets can be sent to D. Due to the intermittent connections between nodes in DTNs, a node may not be able to send all packets to the destination at one encountering. Then, it may take a long time for the remaining packets to be sent to the destination node later on. Thus, a long communication time between a relay node and the destination node can increase the packet transmission throughput between a source node and a destination node. Once the node moves into the transmission range of the destination node, it continuously transmits the packet until the connection breaks.

3.2 The design of the DIG scheme

Since a relay node knows its current location and moving direction and the location of the destination node, the relative angle between itself to the destination node can be calculated. We use d to denote the distance between a relay node carrying packets and the destination. As shown in Figure 2, the DIG scheme consists of three routing phases: (1) initial routing, (2) macro-control routing ($d > T$) and (3) micro-control routing ($r < d \leq T$). A small T leads a long macro-control delay whereas a large T leads to a long micro-control delay.

3.2.1 Initial routing phase

A *promising node* refers to a node that has a high probability to have a long communication time period with the destination node. In the initial routing phase, the source node separates the packet stream into several segments and transmit the segments to its neighbour nodes that are located closer to the destination node regardless of their routing utilities in order to increase the probability that the packets are carried or will be carried by promising nodes moving to the destination.

As shown in Figure 2(a), considering the line connecting a source node and a destination node (SD), let θ be the slope of the line SD and d be the distance of the SD line. When a source node wants to send packets to a destination node, it selects neighbours that it meets to forward the packet segments. The selected neighbours must be closer to the destination node than itself, but there is no specific requirement on their moving direction in this initial routing phase. Since the neighbours forward the packets to the destination

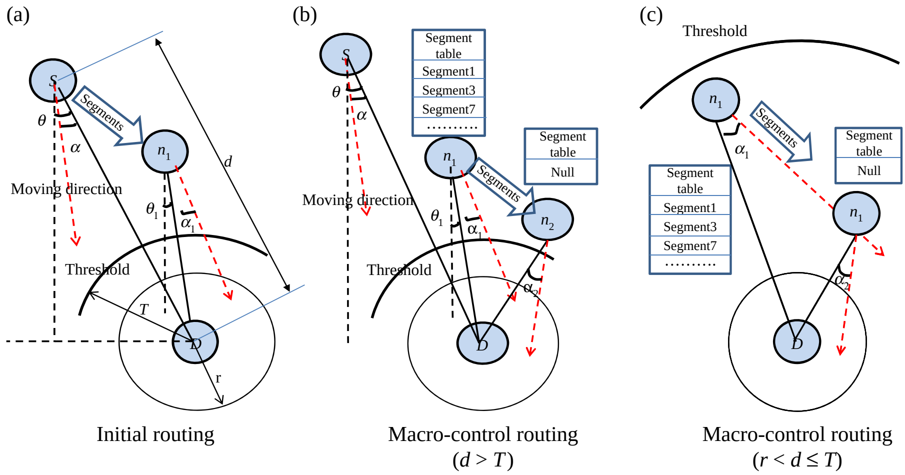

Figure 2. The phases of packet routing in DIG.

node individually, this initial phase aims to forward the packets to the neighbours as more as possible, so that the packets can be forwarded to the destination more quickly. Therefore, node utility is not considered in this phase. A source keeps on forwarding packets to its meeting nodes that are closer to the destination until it sends out all packets. Although some of the nodes selected for packet forwarding may not have high routing utility, the packets will be quickly forwarded to high-utility relay nodes. As these different packet segments can be forwarded towards the destination node concurrently, the transmission delay can be reduced.

3.2.2 *Macro-control phase*

After the packets are forwarded from the source node to the relay nodes, the DIG scheme goes to the second phase. In the macro-control phase, the packets are forwarded to mobile nodes that are close and moving towards the destination node. Macro-control routing aims to quickly forward packets towards the destination node when the packets are far away from the destination. We do not use the routing utility in this phase because a node currently has small relative angle to the destination node or long predicted transmission time may have large relative angle or short predicted transmission time after it changes its moving direction due to the i.i.d. feature of node moving directions.

As shown in Figure 2(b), when $d > T$, relay node n_1 seeks to find another relay node n_2 whose position is closer to the destination node and its moving direction α_2 satisfies

$$\alpha_2 \in [\theta_1 - \xi, \theta_1 + \xi],$$

and

$$\xi = \tau^k \arcsin(r/d) \leq \pi/2(\tau > 1 \ \text{and} \ k \in (0,1)),$$

where r is the transmission range of a node, τ is a weight which increases with time, k is a constant value which controls the changing speed of the angle and θ_1 is the slope of the line connecting node n_1 and destination node. Smaller relative angle (i.e. α) of a node indicates how fast the node moves towards the destination node. Therefore, we first set

ξ to 0 in order to let the current node to select next hop node that moves more closely to the destination. To deal with the situation in which no node satisfies the strict angle constraint, the angle ξ increases over time if a node cannot find a qualified next hop node for packet forwarding. A larger angle enables the node to have more choices for the next hop selection. ξ cannot be more than $(\pi/2)$ because the next hop node must move towards the destination node regardless of the closeness of its moving direction to the destination. Every time after a relay node successfully finds a next hop for packet transmission, τ is reset to 1 and then $\xi = \arcsin(r/d)$ accordingly. We call this phase 'macro-control' phase, because the moving direction constraint is gradually loosed in order to have more choices on the node selection. In this phase, as the relay nodes move closer to the destination node, the packets can be transmitted comparatively faster to the destination node.

3.2.3 Micro-control phase

In DTNs, the communication time between nodes is limited, and the frequency of a node meeting a destination is not high. Therefore, after a relay node passes the destination node, it may take a long time for the node to find another promising relay node to forward the remaining packets to the destination node. The micro-control phase aims to guarantee the communication time between the relay nodes and the destination node as long as possible. As shown in Figure 2(c), when $r < d \leq T$, the DIG scheme goes to the third phase called micro-control phase. In this phase, the packets are further forwarded to the relay nodes with longer expected transmission time (smaller relative angles or longer predicted communication time) to the destination regardless of the location of the nodes. When the relay node is out of the transmission range of the destination, if there are no nodes better utilities (i.e. longer expected transmission time) nearby, the relay node will continually carry the packets by itself until it reaches the destination node or meets a better node. Once $d < r$, the node continuously transmit the packets to the destination node until the connection breaks. We introduce the two utilities in the next section.

3.3 *Transmission time utilities*

3.3.1 Relative angle utility-based relay selection

When two nodes stay close to each other, the node with closer transmission direction to the destination node has more communication time for the packet transmission to the destination. Then, smaller relative angle (i.e. α) of a node indicates longer communication time between the node and the destination. However, when two nodes are far away to each other, the relative angle α cannot accurately reflect the communication time between the mobile node and a destination node. Figure 3 shows that although $\alpha_0 < \alpha_3$, node n_3 will move close to the destination faster than n_1. Only when two nodes are in the transmission range of each other, they need to compare their relative angle utilities to determine which node is a more promising node. Therefore, only when transmission ranges of two nodes are small, the relative angle can accurately reflect the communication time.

The relative angle calculation in relay node selection does not need much computation and consumes less energy. Therefore, the relative angle-based relay node selection algorithm is easy to implement in mobile networks. When the transmission range is small, DIG uses the relative angle utility. That is, in the micro-control phase shown in Figure 2(c), a relay node n_1 forwards packets to another node n_2 whose moving direction (α_2) to the destination node is smaller than itself (α_1), even if the position of node n_2 is further to the destination node.

Algorithm 1 Pseudo-code of the DIG packet routing scheme executed by node n_j.

```
while n_j.packets.size > 0 do
  //sort the packets according to the priority and time_stamp
  n_j.sortPackets();
  for every packet pkt_i do
    dest = pkt_i.getDest();
    //the destination node is in the transmission rage of n_j.
    if n_j.withinRange(dest) then
      //forward packet i to the destination node
      n_j.transmit(pkt_i,dest);
    else
      for every node n_k within the transmission range of n_j do
        //insert the position and moving direction information of n_k into n_j's routing table
        location = n_k.getLoc();
        move_d = n_k.getMovingDirection();
        n_j.routingTable.insert(n_k, location, move_d);
        if n_j == pkt_i.getSource() then
          //initial routing phase when n_j is the source node of pkt_i
          if distance(n_k, dest) < distance(n_j, dest) then
            n_j.transmit(pkt_i, n_k);
          end if
        else if distance(n_j, dest) > T then
          //macro-control routing phase when the distance between pkt_i's dest is beyond threshold T
          if distance(n_k, dest) < distance(n_j, dest) && n_k.angleUtility.between(θ − ξ, θ + ξ) then
            //the n_k is closer to the destination and n_k's angle utility is between [θ − ξ, θ + ξ].
            n_j.transmit(pkt_i, n_k);
          end if
        else if distance(n_j, dest) ≤ T then
          //micro-control routing phase when the distance between pkt_i's dest is within threshold T
          if relative angle utility is used to predict the communication time then
            if n_k.angleUtility > n_j.angleUtility then
              n_j.transmit(pkt_i, n_k);
            end if
          else
            //communication time is used as the utility in the micro-control routing phase
            if n_k.timeUtility > n_j.timeUtility then
              n_j.transmit(pkt_i, n_k);
            end if
          end if
        end if
      end for
    end if
  end for
end while
```

3.3.2 Predicted communication time-based relay selection

Because of the limitations of the relative angle-based routing, a more accurate transmission time prediction method is needed when the transmission range of mobile nodes is large. Below, we introduce how to directly calculate the communication time between a node and a destination given the node's current location and moving direction, and the destination's location. The predicted communication time is a more accurate routing utility than the relative angle.

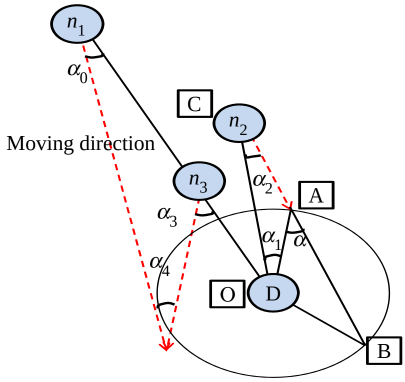

Figure 3. Calculation of the communication time.

In Figure 3, we use v to denote the moving speed of a mobile node, and L to denote the length from point A to point B, represented by $AB = L$. Then, L/v is the predicted communication time with a destination node of a mobile node moving from A to B. Similarly, we define the following notation: $AO = BO = r$, $OC = d$, $\angle OAB = \alpha$, $\angle OCB = \alpha_2$, $\angle AOC = \alpha_1$. As Figure 3 shows, $\alpha = \alpha_2 + \alpha_1$. Therefore

$$\sin(\alpha) = \frac{d}{r} \cdot \sin(\alpha_2) = \frac{\sqrt{r^2 - (L/2)^2}}{r},$$

$$\Rightarrow L = 2\sqrt{r^2 - d^2 \sin^2(\alpha_2)},$$

$$\Rightarrow T_{\text{comm}} = \frac{2\sqrt{r^2 - d^2 \sin^2(\alpha_2)}}{v},$$

where T_{comm} denotes the predicted communication time of a node with a destination node. Although the routing utility of the predicted communication time has higher prediction accuracy than the relative angle-based prediction when the transmission range of mobile nodes is large, it generates more computation overhead. Therefore, for mobile nodes with a short transmission range, the relative angle based routing is still a better method.

3.4 Packet handoff management

To ensure a short average transmission delay, a packet should be delivered to the destination node as soon as possible. A relay node may have several packets with various destinations at a time. However, the number of packets that can be transmitted during each communication is limited due to the limited communication time. The transmission link between two nodes breaks if one node moves out the transmission range of the other node. DIG incorporates a packet handoff management scheme to decide which packets have higher priority to be transmitted during the communication time in order to achieve a short average transmission delay in the system.

In each packer's head, in addition to the fields such as the IDs for the source and destination node, DIG includes two new fields: *priority* and *timestamp*. *Priority* is used to indicate the delivery urgency of packets indicated by the applications. *Timestamp* is used to record the elapsed time since a packet's creation. In a node's buffer, the packets are arranged in decreasing order of *priority*. Within each level of priority, the packets are sorted in decreasing order of *timestamp*. When two nodes meet each other, the packets are delivered based on the ordered sequence in the buffer. Employing *timestamp* guarantees that the packets staying in a buffer for a longer time have higher priority to be delivered out. Thus, DIG avoids the situation in which a packet stays in a buffer for a very long time period.

Algorithm 1 shows the pseudocode of the DIG packet routing scheme executed by a node. First, node n_i sorted its packets according to the *priority* and $time_s tamp$ (lines 2 and 3). For each packet, if n_i is in the transmission range of the packet's destination, it directly forwards the packet to the destination (lines 6–9). Otherwise, n_i executes one of the three phases for the packet. Specifically, if the packet is generated by n_i, it executes the initial routing for the packet (lines 16–20). If the distance between n_i and the packet's destination is $>T$, then n_i executes the macro-control routing for the packet (lines 21–26). Otherwise, n_i executes micro-control routing for the packet (lines 27–38). In this phase, n_i use either the relative angle utility (lines 29–32) or the communication time utility (lines 34–37).

4. Theoretical analysis

In delay tolerant networks, since the nodes occasionally meet each other, the transmission interference between nodes is not severe. Therefore, we do not consider collisions and hidden terminal problems that are inherent in MANETs in our analysis.

4.1 Transmission throughput

We refer to a node's encountered nodes that it can deliver its packets to as the node's handoff nodes. We discuss the number of handoff nodes that a mobile node M_i can have in a time period $\bar{t}$ as below. As indicated in the paper [1], a node M_j can be a handoff node of M_i in the time interval $\bar{t}$ if both of the following conditions are satisfied.

(1) M_j moves in the direction $x \in [\theta - \xi, \theta + \xi]$, where θ is the relative angle between M_i and M_j and ξ is an angle constrain for the moving direction of node M_i.
(2) M_j encounters node M_i at some time during $[t_0, t_0 + \bar{t}]$, where t_0 is the start time.

We use m to denote the number of nodes that the source node M_i meets during $\bar{t}$ and use v to denote the average moving speed of a node.

Theorem 4.1. Let M_i be a mobile node seeking for handoff nodes and Y_j be a Bernoulli indicator random variable, where $Y_j = 1$ if a node M_j is a handoff node of M_i. For any $\theta \in [0, 2\pi)$, the average number of handoff nodes of M_i during a time period $\bar{t}$ depends on the node density and moving speeds of mobile nodes. Particularly, it satisfies:

$$\sum_{i=0}^{m-1} E[Y_j] > 2m\bar{t}^2|v|^2 \cdot (\theta + \xi - \sin(\theta + \xi)).$$

Proof. Based on conditions (1) and (2), the number of handoff nodes of M_i during a time period $\bar{t}$ equals the average number of handoff nodes that M_i meets within the angle range $x \in [\theta - \xi, \theta + \xi]$ in $\bar{t}$. That is,

$$E[Y_j] = \int_{\theta-\xi}^{\theta+\xi} E[Y_j | d_i = x] P[d_i = x] \mathrm{d}x,$$

where d_i is a variable that indicates an angle and P is probability that M_j stays at direction x. The relative speed of M_i to M_j is $|2|v|\sin(x/2)|$. Thus, given that $d_i = x$, M_i can meet a node M_j iff M_i lies in a region of area

$$E[Y_j|d_i = x] = \pi L^2 = \pi \cdot \bar{t}^2|2 \cdot |v| \cdot \sin\frac{x}{2}|^2$$

,where L is the distance between two nodes that may depart from each other during time interval $\bar{t}$.

$$E[Y_j] = \frac{1}{2\pi}\int_{\theta-\xi}^{\theta+\xi} \pi L^2 \cdot \mathrm{d}x = \frac{1}{2\pi}\int_{\theta-\xi}^{\theta+\xi} \pi\bar{t}^2|2 \cdot |v| \cdot \sin\frac{x}{2}|^2 \cdot \mathrm{d}x.$$

$$E[Y_j] = \frac{1}{2\pi} \cdot \pi\bar{t}^2 \cdot 4|v|^2 \cdot 2\int_0^{\theta+\xi} \sin\frac{x^2}{2} \cdot \mathrm{d}x,$$

$$= 2\bar{t}^2|v|^2(\theta + \xi - \sin(\theta + \xi)).$$

Then,

$$\sum_{i=0}^{m-1} E[Y_j] > 2m\bar{t}^2|v|^2 \cdot (\theta + \xi - \sin(\theta + \xi)).$$

□

The more handoff nodes, node M_j meets within time $\bar{t}$, the higher throughput node M_1 has. Therefore, based on Theorem 4.1, the throughput of a node can be determined based on node density and mobility.

Theorem 4.2. Let λ denotes the data transmission rate of a node, then the amount of data that a mobile node can handoff to other nodes during the time period $[t_0, t_0 + \bar{t}]$ is at least $\lambda r_0 m\bar{t}^2|v|O(\sigma)$.

Proof. We use $|v|$ to denote the absolute velocity of each of the mobile nodes. Considering the motion of a mobile node M_j relative to a mobile node M_i, the speed of M_i relative to M_j (i.e. relative speed) is at most $2|v|$ if they move in opposite directions. Moreover, if M_j moves a distance of r_0 during its encountering time with M_i, the time duration of the transmission is $(r_0/2|v)$. Hence, the expected amount of data delivered during an encounter is at least $(\lambda r_0/2|v|)$. Followed by the Theorem 4.1, the amount of data that a mobile node can handoff to other nodes in $[t_0, t_0 + \bar{t}]$ is at least $\lambda r_0 m\bar{t}^2|v|o(\sigma)$. □

4.2 Transmission bound of DIG

Recall that in the UMM, mobile nodes are uniformly distributed initially, moving at a constant speed v, and the moving directions are i.i.d in the range $[0, 2\pi]$. We assume that the network size of DTN is fixed. In order to make the analysis clear, as shown in Figure 4, we use a big unite square to represent the entire network area and partition it into several small cells with a cells in each side. The transmission range of a node can cover the entire cell it is located. Therefore, the average length of each cells is $1/a$. A node in each cell can only communicate with the nodes in its neighbouring cell. There are total n number

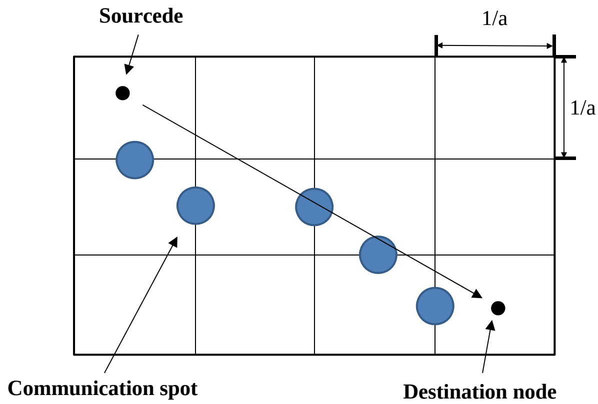

Figure 4. Communication routing path in DIG.

of mobile nodes, among which N_s nodes are randomly chosen as senders, whereas the rest of nodes act like possible packet receivers [12].

Proposition 4.3. The transmission delay in DIG is lower bounded by $O(n/a)$.

Proof. The average delay of DIG includes the inter-cell transmission delay from the source cell to the destination cell and the intra-cell transmission delay for locating destination in a cell. The former is the product of the number of transmission hops, $O(\bar{L}/(1/a))$ and the delay in each hop $O(1/(va))$. The speed of the nodes is $v = O(1/\sqrt{n}))$ [7]. Then, the average inter-cell transmission delay is

$$\frac{O(\bar{L})}{1/a}\frac{1}{v \cdot a} = O(\sqrt{n}).$$

If the destination node did not move out of the destination cell when the message arrives at the cell, the intra-cell transmission delay is $O(\sqrt{n})$. Otherwise, the intra-cell transmission delay is $O(\sqrt{n} + n/a)$, where $O(n/a)$ is the lower bound for locating the mobile destination node in the network [12]. Since $O(n/a) \ll O(\sqrt{n})$, the total delay is $O(n/a) + O(\sqrt{n}) \approx O(n/a)$, which is nearly the same as single-copy opportunistic routing scheme. That is, the performance of DIG is lower bounded by the single-copy opportunistic routing when the destinations are mobile nodes [12]. □

4.3 Discussions on threshold T

To calculate the packet delivery delay between a source node and a destination node, we need to calculate the delay in the three packet forwarding steps. Suppose the distance between a source node and a destination node is d, as the nodes are i.i.d in the networks based on our assumption, the probability for the source node to meet a node that is located closer to the destination node in the first step is $1/2$. Suppose the average delay for a node meets another node is $\bar{t}$, the average delay for the source node to meet a node that is located closer to the destination node should be $2 \cdot \bar{t}$. After the first step, the expected distance between the forwarding node and the destination node is $d - r - \bar{t} \cdot \bar{v}$. In the second step, the probability of a node meets a node that is closer and moves towards the destination node is $(1/2)\cdot(1/2) = 1/4$. Then the expected delay of a forwarding node meets a node that is closer and moves towards the destination node is $4\cdot\bar{t}$. Then the average delay in the second step is $(4\bar{t}(d - T - r - \bar{t}\bar{v}))/(2\bar{t}\bar{v} + r)$. In the third step, after a node moves beyond threshold T, the node only forwards the packets to the nodes that have closer direction to the destination node. Therefore, the expected delay for a node meets a next hop node that meets the requirement is $(\pi/\alpha)\cdot\bar{t}$, where α is the direction angle between forwarding node and the destination node. The expected distance the forwarding node moves before meeting a satisfied node is $(\pi/\alpha)\cdot\bar{t}\cdot(\bar{v}/2)$. If $(\pi/\alpha)\cdot\bar{t}\cdot(\bar{v}/2) < T - r$, the packet is forwarded to another node with α_i, which is less than α. Therefore, the third step can be approximated to $(T/\delta)\cdot(\pi/\alpha)\cdot\bar{t}$, where $(T/\delta) > 1$. Therefore, the expected overall packet transmission delay is

$$T_{\text{trans}} = 2\bar{t} + \frac{4\bar{t}(d - T - r - \bar{t}\bar{v})}{2\bar{t}\bar{v} + r} + \frac{T}{\delta}\cdot\frac{\pi}{\alpha}\cdot\bar{t}.$$

$$\Rightarrow T_{\text{trans}} = 2\bar{t} + \frac{4\bar{t}(d - r - \bar{t}\bar{v})}{2\bar{t}\bar{v} + r} + \left(\frac{\pi\bar{t}}{\alpha\delta} - \frac{4\bar{t}}{2\bar{t}\bar{v} + r}\right)T.$$

Therefore, our objectives are listed in the formulated problem below.

(1) Minimize : $$T_{\text{trans}} = 2\bar{t} + \frac{4\bar{t}(d - r - \bar{t}\bar{v})}{2\bar{t}\bar{v} + r} + \left(\frac{\pi\bar{t}}{\alpha\delta} - \frac{4\bar{t}}{2\bar{t}\bar{v} + r}\right)T.$$

(2) Maximise : $$T_{\text{comm}} = \frac{2\sqrt{r^2 - d^2\sin^2(\alpha_2)}}{v}.$$

When $(\pi\bar{t}/\alpha\delta) - (4\bar{t}/(2\bar{t}\bar{v} + r)) > 0$, a large T leads to a small α, resulting in a larger communication time with the destination node according to Equation (1). However, a small T leads to a small packet deliver delay as shown in Equation (2). As the packet throughput is determined by both communication time and deliver delay, there is a tradeoff in T. We will numerically evaluate the tradeoff value of T in Section 5.

5. Performance evaluation

This section demonstrates the distinguishing properties of DIG through simulations on OMNeT++ [31]. We used the epidemic routing scheme [43] to represent the flooding-based schemes, and the direct transmission scheme [12] to represent the single-copy opportunistic routing schemes. In the figures of the experimental results, Epidemic denotes the former and Direct denotes the latter. DIG-time represents the DIG routing scheme using the predicated communication time as the routing utility and DIG-angle represents the DIG routing scheme using the relative angle as the routing utility in the micro-control routing phase of DIG.

We conducted simulations based on two models: UMM and Human Mobility Model (HMM) based on the trace data set from the MIT Reality mining project [42]. In UMM, 50 nodes were initially i.i.d. placed in a network with 1500 m × 1300 m space area. The nodes moved at a speed randomly selected from 0 to 20 m/s. In HMM, 94 nodes represent 94 students and staff at MIT. We converted the records of the connections with cellular towers in the real trace to infer each node's mobility for the simulation.

In both models, three nodes were randomly chosen to be static nodes serving as destination nodes. The locations of the destination nodes were known to all the nodes in the system. Each of the other nodes generated one packet at every 5 s for 2000 s and sent the packet to a destination node that is randomly chosen from the three destination nodes. The simulation then ran for another 2000 s to enable the packets to have enough time to be delivered to the destinations in the DTN. The distance threshold T in DIG is set to $2r$, where r denotes the transmission range of mobile nodes. We conducted experiments in two cases: $r = 50$ m and $r = 100$ m in order to see the impact of the transmission range on the routing performance. The TTL in the epidemic routing scheme was set to five hops. Unless otherwise specified, every node has a buffer with a size of 100 packets. A node needs to drop some packets if the number of total packets exceeds the buffer size. In the simulations, we varied the buffer size of each node from 10 to 200 packets. Dropped packets were not to be retransmitted again. Three simulation metrics were used in the simulation:

(1) *Packet delivery delay.* This is the average latency in seconds for a packet to be delivered to its destination. We do not consider unsuccessfully delivered packets. This metric represents the efficiency of a routing scheme.
(2) *Number of successfully delivered packets.* This is the number of packets that are successfully delivered to their destinations. This metric represents the robustness and delivery capacity of a routing scheme.
(3) *Number of forwardings.* Forwarding occurs when a node forwards a packet to another node in the routing. This metric reflects the transmission overhead and resource consumption of a routing scheme.

5.1 Packet delivery delay versus buffer size

Figure 5 plots the packet delivery delay versus buffer size with transmission range equals to 50 and 100 m, respectively, based on the UMM model. Both of the figures show that Direct generates much longer delay than others. This is because, in Direct, the packet transmission delay is based on the meeting probability of the relay nodes and the destination nodes. Since only a single copy of a packet is transmitted in the network and the meeting probability between packet relay nodes and destination nodes is small, the transmission delay in Direct is long. In contrast, relying on flooding, Epidemic uses all possible routing paths to the destination, resulting in a low transmission delay in both figures.

Figure 5(a) shows that when the buffer size is small, the packet delivery delay of Epidemic is slightly smaller than that of DIG. Epidemic generates more packets in transmission than DIG. Thus, Epidemic drops more packets than DIG when the node buffer size is small. The dropped packets are not counted in packet delay calculation.

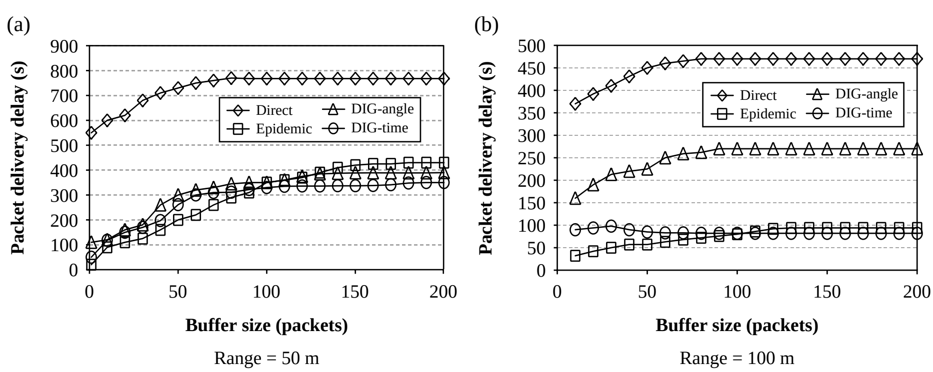

Figure 5. Packet delivery delay versus buffer size in UMM.

The packets that are not dropped can be quickly transmitted to the destination due to the packet flooding. As a result, the packet delivery delay of Epidemic is smaller than that of DIG. However, as the node buffer size increases, more packets can be stored in the buffer which leads to a high packet queuing delay in Epidemic. Therefore, the overall packet delivery delay in Epidemic increases. Since packet queuing delay experienced by DIG is much less than that by Epidemic, the packet delivery delay of DIG is shorter than that of Epidemic with large buffer size. Although DIG also uses a single copy of packets in routing, instead of waiting for the relay nodes to meet a destination node by chance, DIG routes the packets to the destination nodes based on the location information of nodes. Thus, DIG reduces the transmission delay of Direct significantly. Furthermore, the figures show that the delay of DIG is not as sensitive to the buffer size as Epidemic. DIG uses single-copy transmission. Therefore, its total number of packets is almost constant if the buffer is large enough to store all packets, which leads to a constant packet queuing delay. While in Epidemic, the number of packets in buffer increases as the buffer size increases because of the flooding. As a result, the packets experience much longer packet queuing delay.

Figure 5(b) shows that the packet delivery delay of DIG-time is comparable to Epidemic, when the packet transmission range is large. Also, DIG-time has even shorter delay than Epidemic when the buffer size is large. The reason is the same as in Figure 5(a). However, we can see from the figure that DIG-angle generates much longer transmission delay than DIG-time. This is because, with long transmission range, the nodes can communicate with each other within a long distance with long transmission range. In this case, a relative angle between two nodes cannot accurately reflect the communication time between two nodes. Therefore, in DIG-angle, sometimes the packets may be forwarded to the relay nodes that are not moving towards the destination node or have short communication time with the destination node. As a result, the transmission delay of DIG-angle is longer than DIG-time.

Comparing Figure 5(a),(b), we find that as the transmission range increases, the transmission delay of all routing schemes decreases. Intuitively, larger transmission range makes it easier for a node to find more neighbour nodes, which may either be the destination nodes or promising relay nodes, thus leading to a shorter delay.

Figure 6 plots the packet delivery delay versus buffer size with transmission range equals to 50 and 100 m respectively based on the HMM model. As shown in Figure 6(a), when nodes have 50 m transmission range, Direct has the highest packet delivery delay. When the buffer size is small, the packet delivery delay of Epidemic is smaller than DIG. However, as the node buffer size increases, the packet delivery delay in Epidemic increases. The delay of Epidemic is larger than DIG when buffer size is larger than 100 packets. These results are caused by the same reasons as in Figure 5(a).

Figure 6(b) shows the packet delivery delay when nodes have 100 m transmission range. We see that the packet delivery delay of DIG-time is comparable to that of Epidemic and DIG-angle has much longer transmission delay than DIG-time. Direct still has the highest delivery delay. The observations are consistent with Figure 5(b) due to the same reasons.

Comparing Figures 5 and 6, we find that all of the protocols in HMM (Figure 6) have much less delay than in UMM (Figure 5) except DIG-angle when nodes have long transmission range. The reason is that the node density in HMM is much higher than that in UMM. Then, the nodes have much higher interaction frequency between each other in HMM than in UMM, and hence the transmission delay in HMM is less than that in UMM. In DIG-angle, the nodes cannot predict the communication time between the nodes very

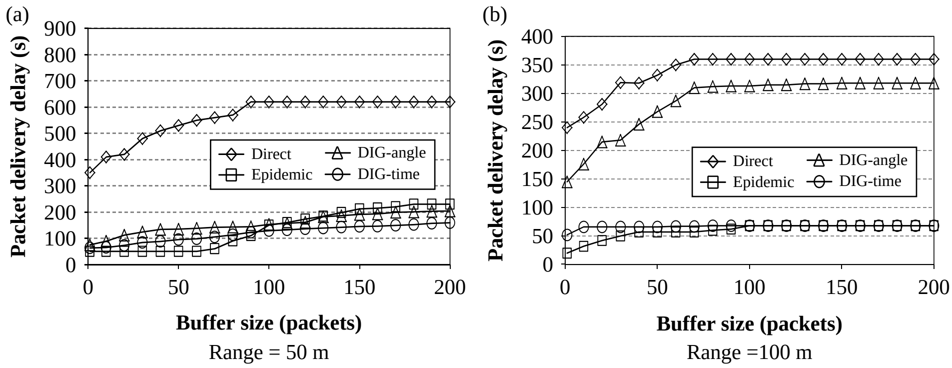

Figure 6. Packet delivery delay versus buffer size in HMM.

accurately. Therefore, when the system has more nodes, it is more likely for a mobile node to mistakenly forward the packets to the next-hop node that has very short communication time with the destination node, which further increases the packet transmission delay.

5.2 *Packet delivery capacity versus buffer size*

Figure 7 depicts the number of successfully delivered packets versus the node buffer size with transmission range equals to 50 and 100 m in UMM model, respectively. It shows that as the buffer size increases, so does the number of the successfully delivered packets. A larger buffer size means more packets can be buffered and less packets are dropped, resulting in more successfully delivered packets. Figure 7 also shows that the number of successfully delivered packets in DIG and Direct increases much faster than Epidemic. By flooding, Epidemic generates many packets, which overwhelm node buffers. Hence, many buffer congestions occur and many packets are dropped. DIG and Direct do not have buffer congestion problem due to their single-copy opportunistic routing, whereas Epidemic with flooding suffers from buffer congestion severely, especially when the buffer size is small. We also find that DIG generates more delivered packets than Direct with a smaller buffer size. This is because the transmission delay of DIG is much shorter than that of Direct. Thus, DIG has much more free buffer space all the time. Therefore, when the buffer size is small, DIG has a smaller chance to experience data congestion than Direct.

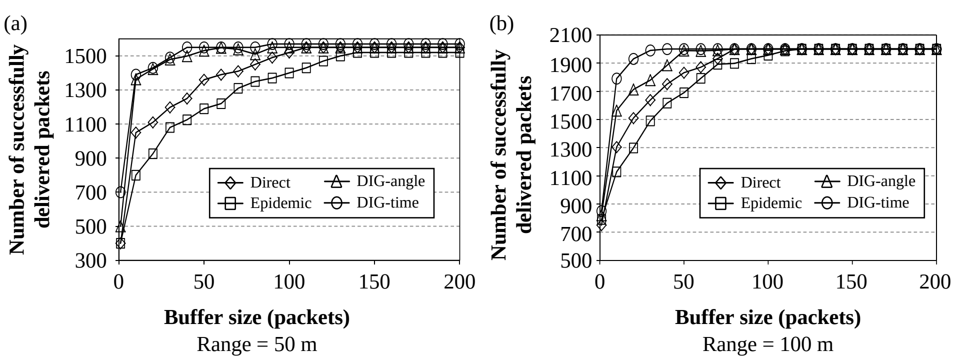

Figure 7. The number of successfully delivered packets versus buffer size in UMM.

Moreover, Figure 7 indicates that when the buffer is large enough for all the packets, the delivery ability of all the routing schemes are almost the same. This is because there is no buffer congestion during the transmissions. DIG performs the best among the schemes with regard to packet delivery ability when the buffer size is small. This is caused by two reasons. First, compared to Epidemic, since DIG only uses one copy of packets in routing, the performance of DIG is not greatly affected by the buffer size. Second, compared to Direct, DIG uses the location and moving direction information of mobile nodes to quickly forward packets to the destination nodes and to ensure a long communication time between relay nodes and the destination nodes.

Comparing Figure 7(a),(b), we can observe that as the transmission range increases, the number of successfully delivered packets increases. The result is consistent with Theorem 4.1, which shows that with the increase in the transmission range, the communication time between two mobile nodes increases. Also, as the probability of a mobile node meeting the destination node or promising forwarding nodes increases as transmission range increases, the number of successfully delivered packets increases. It is intriguing to see that when the transmission range of the mobile nodes is small, DIG-angle and DIG-time produce nearly the same number of successfully delivered packets. But when the transmission range is large, DIG-time has a much higher packet delivery capacity than DIG-angle. This is because DIG-angle cannot accurately predict the communication time between a relay node with the destination node when the transmission range is large as we discussed in Section 3.1.

Figure 8 depicts the number of successfully delivered packets versus the node buffer size with transmission range equals to 50 and 100 m in the HMM model, respectively. The figure shows that as the buffer size increases, so does the number of the successfully delivered packets. The figure also shows that the number of successfully delivered packets of DIG and Direct increases much faster than that of Epidemic. Meanwhile, as the transmission range increases, the number of successfully delivered packets increases. The results are in line with Figure 7 due to the same reasons. However, comparing Figures 7 and 8, we can find that the number of successfully delivered packets in Epidemic in UMM increase much faster than that in HMM, while those of other individual corresponding schemes in both figures are almost the same. As there are more nodes in the HMM model, Epidemic generates more copies of packets. Therefore, the nodes are more likely to be congested when the nodes have a small buffer size, which leads to the droppings of a large number of packets. In all other schemes, since there is only one copy of packets that is forwarded by the nodes in the system,

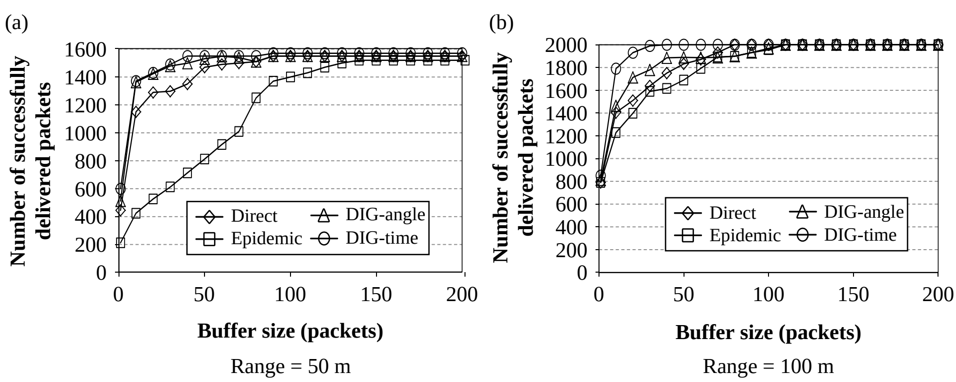

Figure 8. The number of successfully delivered packets versus buffer size in HMM.

the nodes are less likely to be congested. Therefore, the delivery capacity of all other schemes does not change greatly as the number of nodes in the system increases.

5.3 Transmission overhead versus buffer size

Figure 9 shows the number of forwardings versus the buffer size in UMM. The figure shows that DIG and Direct incur much fewer forwardings and hence a much smaller communication overhead than Epidemic (Figure 10). This is due to the reason that Epidemic uses flooding in packet transmission, in which a node sends all packets to all nodes it encounters. In contrast, DIG and Direct only forward one copy of the packets in the network. That is also why Epidemic experiences packet congestions when the mobile nodes have small buffer size. Figure 11 further shows the detailed comparisons of the number of forwardings versus the buffer size in UMM for Direct, DIG-angle and DIG-time. The figure shows that DIG-angle has the largest number of forwardings. This is because the relative angle utility has lower accuracy to reflect the communication duration between a relay node and the destination node than the predicted communication time utility. Since in Direct routing mechanism, a relay node will hold the packets all the way to the destination node rather than forwarding for more hops, Direct has the smallest number of forwardings.

Figure 10 shows the number of forwardings versus the buffer size in HMM. The figure shows that DIG and Direct incur much smaller communication overhead than Epidemic, which is inline with Figure 9. Comparing Figures 9 and 10, we can see that the overhead incurred by the schemes in HMM is higher than UMM especially in Epidemic. This is because HMM has more nodes than UMM, and then incurs more forwardings in a routing. Because only one copy of packets is forwarded in Direct and DIG, they suffer less congestion when the buffer size is small. Figure 12 further shows the detailed comparisons of the number of forwardings versus the buffer size in HMM for Direct, DIG-angle and DIG-time. The figure shows that DIG-angle has the largest number of forwardings whereas Direct has the smallest number of forwardings. The reason is the same as in Figure 11.

5.4 Packet delivery delay versus threshold

Figure 13 shows the packet delivery delay of nodes in Direct, DIG and Epidemic versus DIG's distance threshold T in UMM. We can see from the figures that as the threshold increases, the packet delivery delay of DIG-angle and DIG-time decreases and then

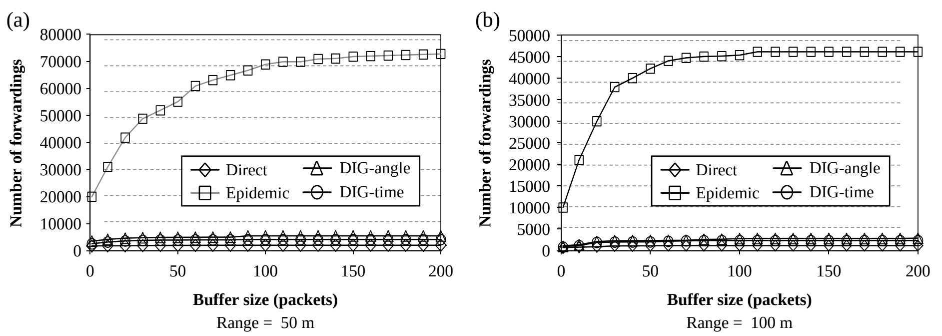

Figure 9. The number of forwardings versus buffer size in UMM.

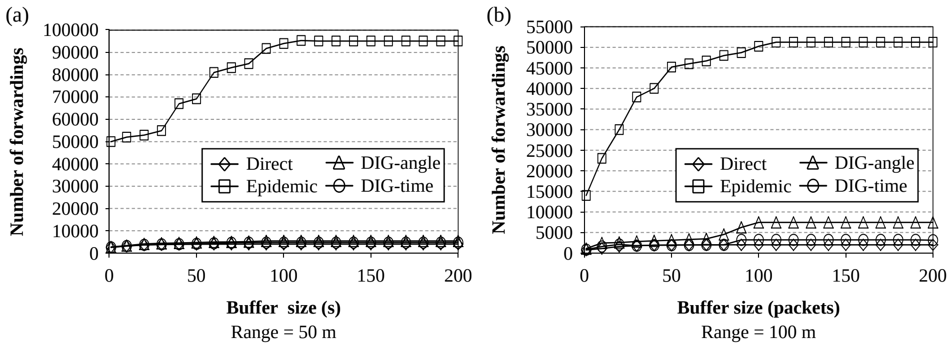

Figure 10. The number of forwardings versus buffer size in HMM.

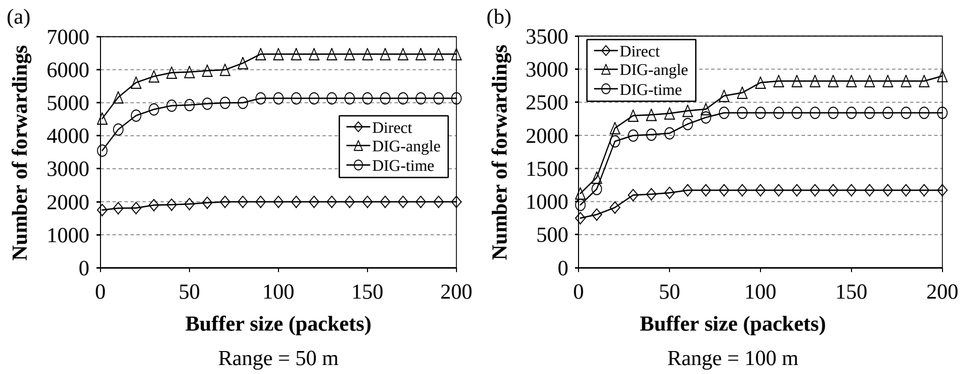

Figure 11. The number of forwardings versus buffer size in UMM.

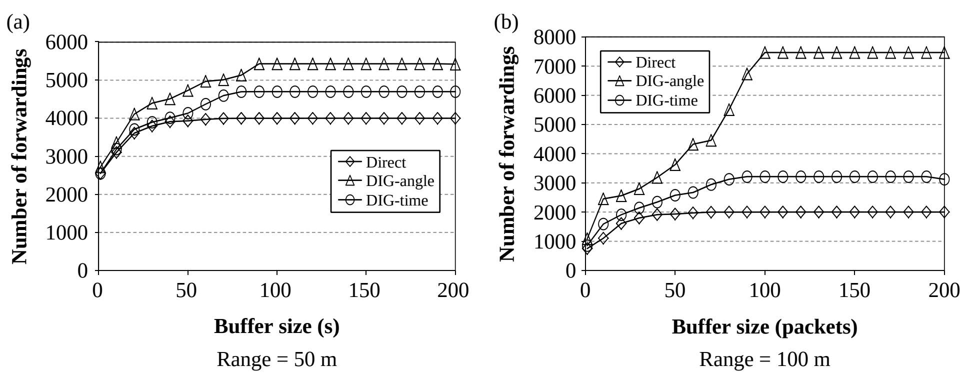

Figure 12. The number of forwardings versus buffer size in HMM.

increases as the threshold increases. It is very interesting to see that in both Figure 13(a),(b), DIG achieves the smallest packet delivery delay when the threshold is almost twice of the transmission range. If the threshold is set too small, although the micro-control time is small, the macro-control time is large, which leads to a large overall packet forwarding delay. In contrast, if the threshold is large, it takes a long time for a

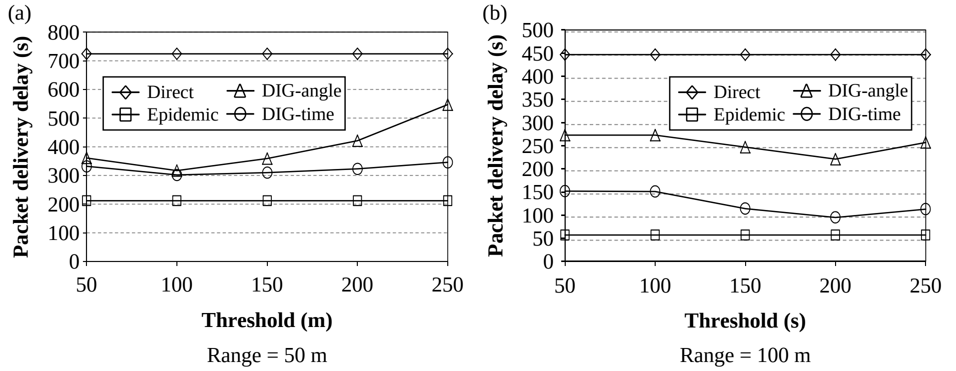

Figure 13. Packet delivery delay versus threshold in UMM.

message to be forwarded to the destination node in the micro-control, even the time consumed in micro-control step is small. The threshold value that is almost twice of the transmission range is the point that achieves the smallest delay. Since there is no threshold constraint in both Direct and Epidemic, their packet delivery delay keeps constant. We can also see from the figure that Direct generates the largest delay. The delay in DIG is larger than that of Epidemic and DIG-angle generates larger delay than DIG-time. The reason is the same as Figure 5.

Figure 14 shows that packet delivery delay of nodes in Direct, DIG and Epidemic versus the distance threshold in HMM. Similar to Figure 13, we can also see from the figure that the smallest packet delivery delay in DIG is achieved when the threshold is almost twice of the transmission range due to the same reasons. Meanwhile, Direct has the largest delay, Epidemic has the smallest delay, and DIG-angle generates larger delay than DIG-time. The results are consistent with Figure 13 because of the same reasons. Comparing Figures 13 and 14, we can see that the nodes in HMM experience much less delay than the nodes in UMM. Higher node density in HMM enables a node to meet nodes with much high frequency. Therefore, the packet can be forwarded towards the destination node with much smaller delay in HMM.

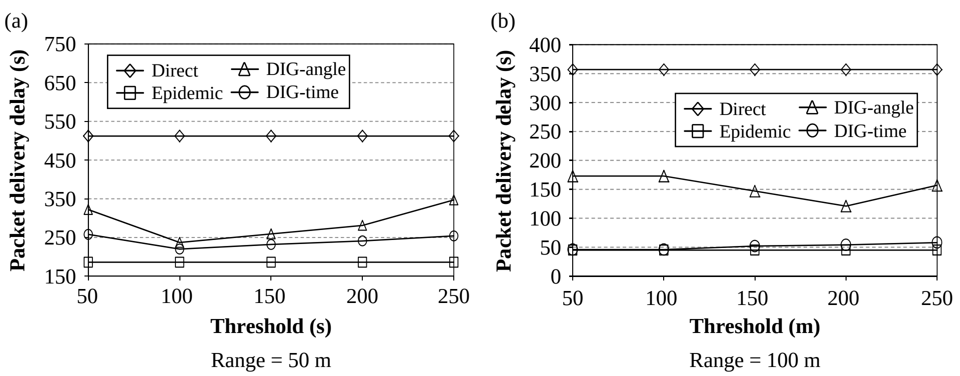

Figure 14. Packet delivery delay versus threshold in HMM.

5.5 Packet delivery capacity versus threshold

Figures 15 and 16 show the number of successfully delivered packets in Direct, DIG and Epidemic versus the distance threshold in UMM and HMM respectively. We can see from the figures that the nodes achieve the largest number of successfully delivered packets when the threshold is the twice of the transmission range of the mobile nodes, whereas the number of successfully delivered packets in Direct and Epidemic keeps almost the same. The reason is the same as Figures 13 and 14.

Comparing Figures 15(a) and 16(a), we can see that when the transmission range of nodes is small, the number of successfully delivered packets of Epidemic in HMM is higher than that of Epidemic in UMM, whereas the number of successfully delivered packet in all other schemes in HMM exhibits a slight increase than that in UMM. Since the node density in HMM is larger than the node density in UMM, the packets can be quickly forwarded to the destination as shown in Figures 13(a) and 14(a). The nodes in Epidemic in HMM experience much less buffer congestion than in UMM, as nodes can quickly forward packets to the destinations with a higher node density, resulting in a significant increase in the number of successfully delivered packets in Epidemic, especially when the buffer size is small. We can also see from the figures that the delivery capacity of the nodes in DIG-angle and DIG-time is almost the same in HMM and UMM. This is because when the transmission range of the nodes is small, the prediction accuracy of the communication time between a node and a destination node are close.

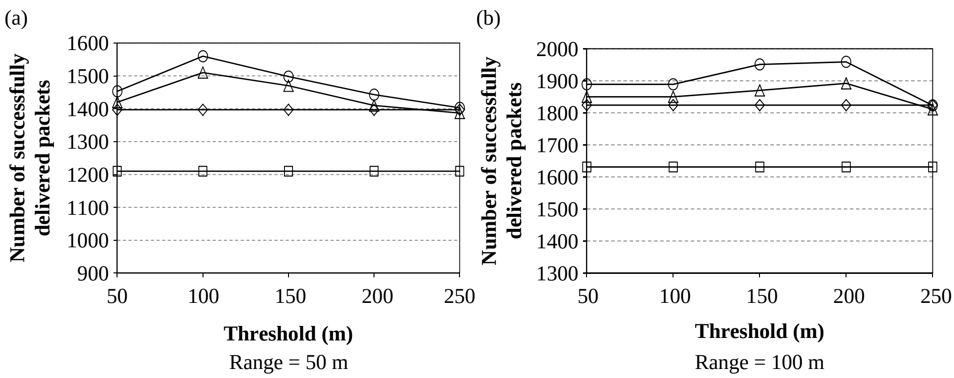

Figure 15. The number of successfully delivered packets versus threshold in UMM.

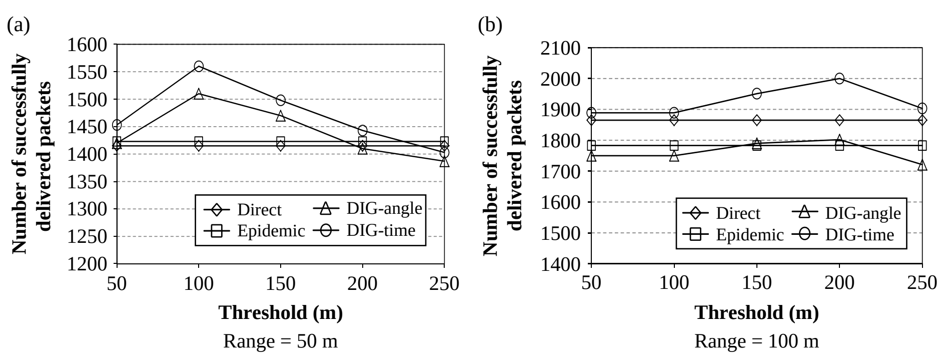

Figure 16. The number of successfully delivered packets versus threshold in HMM.

When the transmission range of the nodes is large, the delivery capability of Epidemic in HMM is much larger than the delivery capability of Epidemic in UMM and the delivery capabilities of the nodes in DIG-time and Direct are slightly increased as shown in Figures 15(b) and 16(b). The reason is the same as Figures 15(a) and 16(a). However, the delivery capability of nodes in DIG-time in HMM is less than the delivery capability of nodes in DIG-time in UMM. This is because DIG-time cannot predict the communication time between a node and a destination node very accurately when the transmission range of the node is very large. As there are more nodes in the system, it is more likely that a node in UMM forwards a packet to another node that does not actually have a long communication time with the destination node.

5.6 *Transmission overhead versus threshold*

Figures 17 and 18 show the number of forwardings of different schemes versus the distance threshold in UMM and HMM. The figure shows that as the threshold increases, the number of forwardings in DIG increases whereas the number of forwardings in Epidemic and Direct keeps almost constant.

This is because a larger threshold leads to longer micro control time. Since the micro control keeps on adjusting the packets among the nodes that potentially have long communication time with the destination node, the overhead in DIG increases. Comparing Figures 17 and 18, we can find that the number of forwardings of all schemes in HMM is

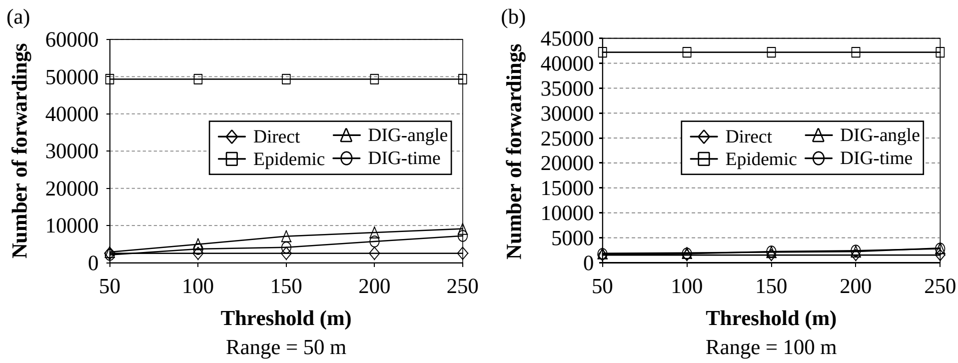

Figure 17. The number of forwardings versus threshold in UMM.

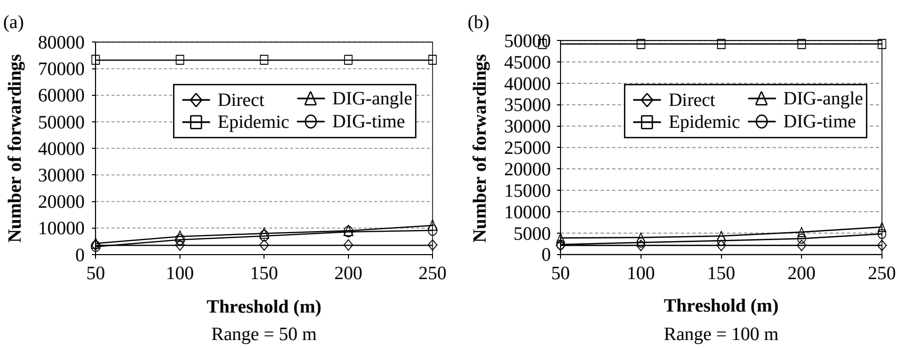

Figure 18. The number of forwardings versus threshold in HMM.

larger than that in UMM. This is because there are more nodes in HMM which may increase forwarding hops between a source node and destination node. The overhead in Epidemic increases most significantly. The messages in Epidemic flood from a source node to a destination node. As there are more nodes in the system, the overhead increases exponentially. Figures 19 and 20 further compare the number of transmission in DIG-angle, DIG-time and Direct, which are extracted from Figures 17 and 18. The figures show that DIG-angle has the largest number of forwardings, Direct has the smallest number of forwardings and the number of forwardings in DIG-time is less than that of DIG-angle. The reason is the same as Figures 11 and 12.

6. Conclusions

In this study, we investigated the problem of efficient routing in intermittently connected mobile networks. Current approaches in such networks are primarily based on redundant transmission or single-copy opportunistic transmission. However, they incur either high overhead due to excessive transmissions or long delay due to incorrect choices during forwarding. We proposed a DIG routing scheme, which overcomes the shortcoming of the redundant transmissions, and reduces the transmission delay of the single-copy opportunistic routing scheme. Using location information that facilitates nodes to be aware of each other's positions and moving directions, DIG tries to forward the packets to the destination node as quickly as possible when the packets are far away from the

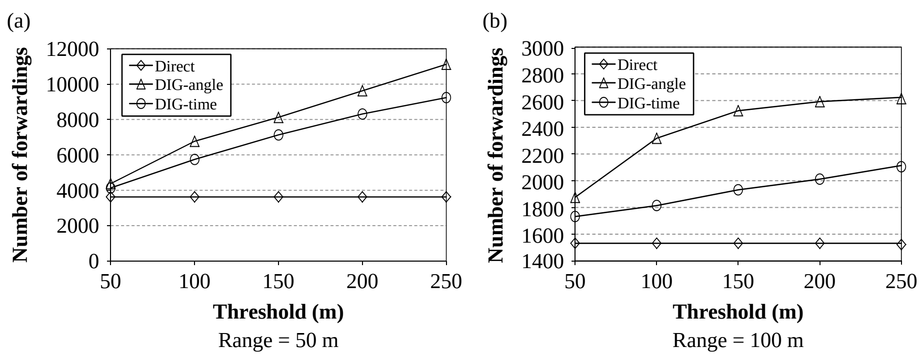

Figure 19. The number of forwardings versus threshold in UMM.

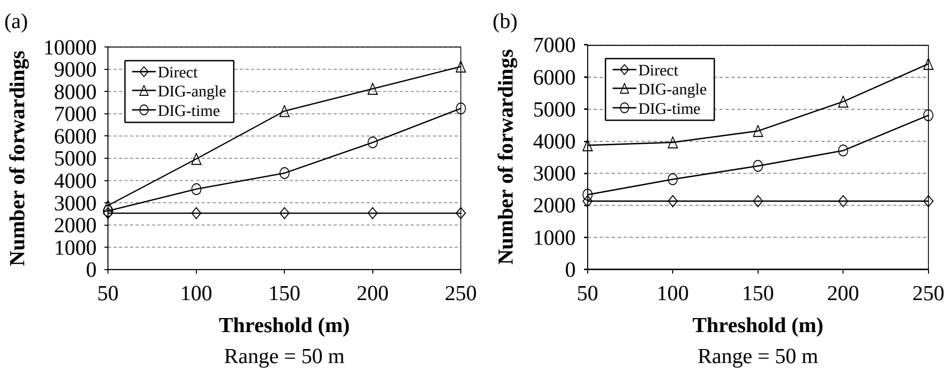

Figure 20. The number of forwardings versus threshold in HMM.

destination node, and tries to forward packets to the relay nodes that have longer communication time with the destination node when the packets are close to it. DIG outperforms the epidemic routing and direct routing with respect to successful transmission, transmission delay and overhead.

Acknowledgements

This research was supported in part by U.S. NSF grants CNS-1249603, OCI-1064230, CNS-1049947, CNS-1156875, CNS-0917056 and CNS-1057530, CNS-1025652, CNS-0938189, CSR-2008826, CSR-2008827, Microsoft Research Faculty Fellowship 8300751 and U.S. Department of Energy's Oak Ridge National Laboratory including the Extreme Scale Systems Center located at ORNL and DoD 4000111689. An early version of this study was presented in the Proceedings of EUC'08 [26].

Note

1. Email: zel@clemson.edu

References

[1] N. Bansal and Z. Liu, *Capacity, delay and mobility in wireless ad-hoc networks*, in *Proceedings of INFOCOM*, 2003.

[2] P. Bose, A. Brodnik, S. Carlsson, E. Demaine, R. Fleischer, A. Lpez-Ortiz, P. Morin, and J. Munro, *Online routing in convex subdivisions*, in *Proceedings of ISAAC*, 2000.

[3] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, *Routing with guaranteed delivery in ad-hoc wireless networks*, ACM Wireless Networks, 2001.

[4] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, B. Durst, V. Cerf, and K. Scott, *Delay-tolerant networking: An approach to interplantary internet*, IEEE Communication Magazine, 2003.

[5] E.P. Charles and P. Bhagwat, *Highly dynamic destination sequenced distance vector routing (DSDV) for mobile computers*, in *Proceedings of SIGCOMM*, 1994.

[6] , Delay tolerant networking research group. http://www.denrg.org

[7] A. El-Gamal, J. Mammen, B. Prabhakar, and D. Shah, *Throughput-delay trade-off in wireless networks*, in *Proceedings of INFOCOM*, 2004.

[8] L. Fang, W. Du, and P. Ning, *A beacon-less location discovery scheme for wireless sensor networks*, in *Proceedings of Infocom*, 2003.

[9] H.D. Ferriere, M. Grossglauser, and M. Vetterli, *Age matters: Efficient route discovery in mobile ad hoc networks using encounter ages*, in *Proceedings of Mobihoc*, 2003.

[10] H. Frey and I. Stojmenovic, *On delivery guarantees of face and combined greedy-face routing in ad hoc and sensor networks*, in *Proceedings of MobiCom*, 2006.

[11] P. Gao, L. Guibas, J. Hershberger, L. Zhang, and A. Zhu, *Online routing in convex subdivisions*, in *Proceedings of Mobihoc*, 2001.

[12] M. Grossglauser and D. Tse, *Mobility increases the capacity of ad hoc wireless networks*, in *TON*, 2002.

[13] T. Hou and V. Li, *Transmission range control in multihop packet radio networks*, IEEE TC, 1986.

[14] S. Jain, K. Fall, and R. Patra, *Routing in a delay tolerant network*, in *Proceedings of ACM SIGCOMM*, 2004.

[15] D.B. Johnson and D.A. Maltz, *Dynamic source routing in ad hoc wireless networks*, IEEE Mobile Computing, 1996.

[16] P. Juang, H. Oki, M. Martonosi, Y. Wang, L.S. Peh, and D. Rubenstein, *Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet*, in *Proceedings of ASPLOS*, 2002.

[17] B. Karp and H. Kang, *Gpsr: Greedy perimeter stateless routing for wireless networks*, in *Proceedings of MobiCom*, 2000.

[18] B. Karp and H. Kung, *Greedy perimeter stateless routing*, in *Proceedings of Mobicom*, 2000.

[19] Y.J. Kim, R. Govindan, B. Karp, and S. Shenker, *Geographic routing made practical*, in *Proceedings of NSDI*, 2003.

[20] E. Kranakis and H. Singh, *Compass routing on geometric networks*, in *Proceedings of CCCGV*, 1999.

[21] F. Kuhn, R. Wattenhofer, and A. Zollinger, *Geometric routing: Of theory and practice*, in *Proceedings of PODC*, 2003.

[22] F. Kuhn, R. Wattenhofer, and A. Zollinger, *Worst-case optimal and average-case efficient geometric ad hoc routing*, in *Proceedings of MobiHoc*, 2003.

[23] B. Leong, B. Liskov, and R. Morris, *Geographic routing without planarization*, in *Proceedings of NSDI*, 2006.

[24] B.N. Levine, A. Balasubramanian, and A. Venkataramani, *DTN Routing as a resource Allocation Problem*, in *Proceedings of SiGCOMM*, 2007.

[25] Q. Li and D. Rus, *Communication in disconnected ad hoc networks using message relay*, ACM/IEEE Transactions on Networking, 2003.

[26] Z. Li and H. Shen, *A direction based geographic routing scheme for intermittently connected mobile networks*, in *Proceedings of EUC*, 2008.

[27] A. Lindgren, A. Doria, and O. Schelen, *Probabilistic routing in intermittently connected networks*, SIGMOBILE Mobile Computing and Communication Review, 2003.

[28] A. Maffei, K. Fall, and D. Chayes, *Ocean instrument internet*, in *Proceedings of AGU*, 2006.

[29] S. Mitra, B. Leong, and B. Liskov, *Path vector face routing: Geographic routing with local face information*, in *Proceedings of ICNP*, 2005.

[30] J.T. Moy, *OSPF: Anatomy of an Internet routing protocol*, Addison-Wesley Professional 1998.

[31] , OMNet++. http://www.omnetpp.org/

[32] J. Partan, J. Kurose, and B.N. Levine, *A survey of practical issues in underwater networks*, in *Proceedings of WUWNet*, 2006.

[33] C. Perkins, E. Belding-Royer, and S. Das, *RFC 3561: Ad hoc on demand distance vector (AODV) routing*,, 2003.

[34] K.S. Pister, L. Doherty, and L.E. Ghaoui, *Convex optimization methods for sensor node position estimation*, in *Proceedings of Infocom*, 2001.

[35] R.C. Shah, S. Roy, S. Jain, and W. Brunette, *Data mules: Modeling and analysis of a three-tier architecture for sparse sensor networks*, Elsevier Ad Hoc Networks Journal, 2003.

[36] H. Shen and Z. Li, *MDR: Market-based distributed routing scheme in hybrid wireless networks*, Technical Report TR-2007-07, CSCE Department, University of Arkansas, 2007.

[37] T. Small and Z. Haas, *Resource and performence tradeoffs in delay-tolerant wireless networks*, in *Proceedings of WDTN*, 2005.

[38] M.A. Sportack and I.P. routing, *IP routing fundamentals*, Cisco Systems 1999.

[39] T. Spyropoulos, K. Psounis, and C. Raghavendra, *Efficient routing in in termittently connected mobile networks: The single-copy case*, in *ACM/IEEE Transactions on Networking*, 2007.

[40] T. Spyropoulos, K. Psounis, and C.S. Raghavendra, *Routing in a delay tolerant network*, in *Proceedings of WDTN*, 2005.

[41] H. Takagi and L. Kleinrock, *Optimal transmission ranges for randomly distributed packet radio terminals*, IEEE TC, 1984.

[42] , The reality mining dataset. http://reality.media.mit.edu/dataset.php

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

[44] Y. Wang, S. Jain, M. Martonosi, and K. Fall, *Erasure-coding based routing for opportunistic networks*, in *Proceedings of SIGCOMM*, 2005.

[45] J. Widmer and J.Y.L. Boudec, *Network coding for efficient communication in extreme networks*, in *Proceedings of WDTN*, 2005.

[46] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, *MDDV: Mobility-centric data dissemination algorithm for vehicular networks*, in *Proceedings of ACM on VANET*, 2004.

[47] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, *Performance modeling of epidemic routing*, in *Proceedings of IFIP*, 2006.

[48] W. Zhao, M. Ammar, and E. Zegura, *A message ferrying approach for data delivery in sparse mobile ad hoc networks*, in *Proceedings of Mobihoc*, 2004.