Volume 5, Issue 2 March 2007 ISSN 1570-8705

Available online at www.sciencedirect.com
ScienceDirect
Ad Hoc Networks 5 (2007) 189–209
Ad Hoc Networks
www.elsevier.com/locate/adhoc

Sociological orbit aware location approximation and routing (SOLAR) in MANET

Joy Ghosh [a,*], Sumesh J. Philip [b], Chunming Qiao [a]

[a] *Department of Computer Science and Engineering, The State University of New York at Buffalo, Buffalo, NY 14260, United States*
[b] *Department of Computer Science, Western Illinois University, Macomb, IL 61455, United States*

Received 4 February 2005; received in revised form 31 August 2005; accepted 31 October 2005
Available online 28 November 2005

Abstract

Mobility affects routing protocol performance in a Mobile Ad Hoc NETwork (MANET). This paper introduces a novel concept of "macro-mobility" information obtained from the sociological movement pattern of MANET users, and proposes a routing protocol that can take advantage of the macro-mobility information. This macro-mobility information is extracted from our observation that the movement of a mobile user exhibits a partially repetitive "orbital" pattern involving a set of "hubs". This partially deterministic movement pattern is not only practical, but also useful for locating nodes without the need for constant tracking and for routing packets to them without flooding.

More specifically, this paper makes the following two contributions. First, it proposes an ORBIT mobility framework to achieve this macro-level abstraction of orbital movement. Second, to take advantage of this hub-based orbital pattern, it proposes a Sociological Orbit aware Location Approximation and Routing (SOLAR) protocol. Extensive performance analysis shows that SOLAR significantly outperforms conventional routing protocols like Dynamic Source Routing (DSR) and Location Aided Routing (LAR) in terms of higher data throughput, lower control overhead, and lower end-to-end delay.

Keywords: Mobility framework; Routing protocol; Ad hoc wireless networks; Performance analysis

1. Introduction

A Mobile Ad Hoc NETwork (MANET) is an infrastructure less group of wireless mobile devices that willfully cooperate to forward packets for one another. Recently, node mobility was shown to have a significant impact on the routing protocol performance [1]. This led to numerous studies on the practicality of the different mobility models that are commonly used in mobile network research. The Random Waypoint model [2] has been a favorite for its simplicity and suitability for theoretical study and analysis. In this model, a node chooses destination points randomly within a terrain and approaches it linearly with a velocity randomly selected from a specified range. In reality however, nodes (i.e., MANET users) move purposefully (e.g., going from a conference room to a cafeteria) while being subject to certain restrictions (e.g., geographical constraints), resulting in certain amount of determinism in their motion.

---

[*] Corresponding author. Tel.: +1 716 308 0082.
E-mail address: joyghosh@cse.buffalo.edu (J. Ghosh).

1570-8705/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2005.10.003

In light of growing need for more practical and realistic mobility modeling, a new line of research has emerged, focusing on several *Entity based* [3–5], *Group based* [6–8], and *Scenario based* [9–11] mobility models/frameworks. *Entity based* models are driven by the individual node characteristics, while *Group based* models concentrate on the collective movement of a group of nodes that deviate marginally from the characteristics of a leader node. Alternately, *Scenario based* models account for the geographical constraints on real life movement.

Literature describes many routing protocols to combat the adverse effect of mobility on MANET users. These can be generally classified into two categories: *Proactive* and *Reactive*. Studies have shown that the latter is more suitable for highly mobile ad hoc networks due to its ability to withstand rapidly changing network topologies. Additional optimizations for reactive protocols include caching of paths [2] to reduce path request overhead, and caching of node velocity [12] to approximate node locations. Alternatively, others [13,14] have suggested the use of virtual backbones to facilitate the adaptation of the routing protocols to mobility. With the advent of newer and affordable technology such as GPS receivers [15], and other localization techniques, location management schemes coupled with routing strategies [16–19] have also shown to offer efficient routing solutions for MANET.

While interference has been shown to significantly affect the capacity of wireless networks in [20–22], the work in [23] exploited node mobility as a type of multi-user diversity, and showed that node mobility may actually help increase the theoretical MANET capacity. In addition, research has also been done [24–26] to explore the use of mobility information obtained via continuous location tracking and "micro-level" mobility prediction of individual nodes, as a potential tool for facilitating next hop selection or link break prediction.

Recent work on delay tolerant networks (DTN) [27] has addressed issues related to routing among nodes (such as satellites and buses) that have deterministic mobility. In addition, routing in intermittently connected MANETs has also been addressed in [28] and most recently in [29]. In particular, the authors in [29] made the observation that a node is likely to come back to a location that it had frequently visited before, and proposed a probabilistic routing scheme whereby each node maintains the so-called "delivery predictability" to each known destination. However, to the best of our knowledge, no prior work has studied the partially deterministic, sociological orbit based macro-mobility pattern of MANET users, and its implications on location management and routing protocols.

In this paper, we propose a macro-level mobility framework called ORBIT, based on a partially deterministic orbital movement pattern of mobile users around some specific places of social interest called *hubs*. The term "macro-level mobility" refers to the fact that our abstraction does not depend on the exact movement within a hub, or in between hubs. Rather, our abstraction only specifies a set of hubs where a node will visit and spend some amount of time, without specifying any rigid schedule or routes (i.e., a partially deterministic movement pattern). We also propose an effective routing scheme for MANET called the sociological orbit aware location approximation and routing (SOLAR) protocol to take advantage of the spatial/temporal locality of the mobile users (nodes) around these hubs. The proposed ORBIT framework is not only general enough to be realistic, but is also specific enough to be useful. In particular, the proposed SOLAR protocol can be practically implemented without a need for constant location updates (or tracking) and flooding. Extensive numerical results are presented to establish the simplicity and superiority of SOLAR over other conventional protocols, such as DSR and LAR in terms of higher data throughput, lower control overhead, and lower end-to-end delay.

The rest of the paper is outlined as follows. In Section 2, we motivate our work by discussing the sociological movement pattern of MANET users, and describe the ORBIT framework. In Section 3, we provide the details of the proposed sociological orbit aware location approximation and routing (SOLAR) protocol and illustrate its use in a common and realistic MANET scenario. In Section 4, we evaluate the performance of SOLAR through simulations, and show that it is superior to DSR and LAR. In Section 5, we present a more detailed description of other related work on both practical mobility models, and mobility sensitive routing protocols to highlight the novelty of our work. We conclude this work in Section 6.

2. Sociological movement pattern

In the real world, users routinely spend a considerable amount of time at a few specific place(s) that we refer to as hub(s). For example, a graduate student in school may visit and spend some significant amount of time in his/her laboratory, a seminar room, or the cafeteria. Although it is hard (or may be even against privacy policies) to keep track of an individual at all times, one can still take advantage of the fact that most users' movements are within and in between a list of hubs. These hubs are usually visited by the individuals in some probabilistic sequence, and constitute a part of the users' *mobility profiles*. For example, even if we do not know the exact location of the graduate student at any given time, given his/her mobility profile we can most probably find him/her in either the laboratory, or the seminar room, or the cafeteria, without having to look all over the building/campus.

This orbital movement pattern is also observed in a time and space based hierarchy. For example, on a typical weekday, the graduate student could leave home for school in the morning, visit the gymnasium in the evening, and return home at night. Similarly, the student may stay in his home town for a few weeks and visit friends and family in other cities over some weekends, forming yet another higher level nation-wide orbit. This hierarchical concept is illustrated in Fig. 1.

Interestingly, an orbit is one of the most natural forms of motion observed in the microscopic world of molecules, as well as in the planetary universe. However, such natural orbits are mostly deterministic, and their continuous motion does not have the notion of special places like hubs.

Note that our orbital movement pattern differs from existing mobility patterns studied in the literature in that it neither models the motion of the users at a micro-level (i.e., on small time scales or within small distances), nor simply predicts user locations via historical/statistical tracking information [24–26]. It also differs from the deterministic mobility patterns assumed within DTN, where either exact locations of a node can be predicted with an appropriate "oracle", or no location information is available. To the best of our knowledge, no prior work has explored the implication of such a macro-level partially deterministic sociological mobility pattern and its application to location approximation and routing in MANET, despite its practicality.

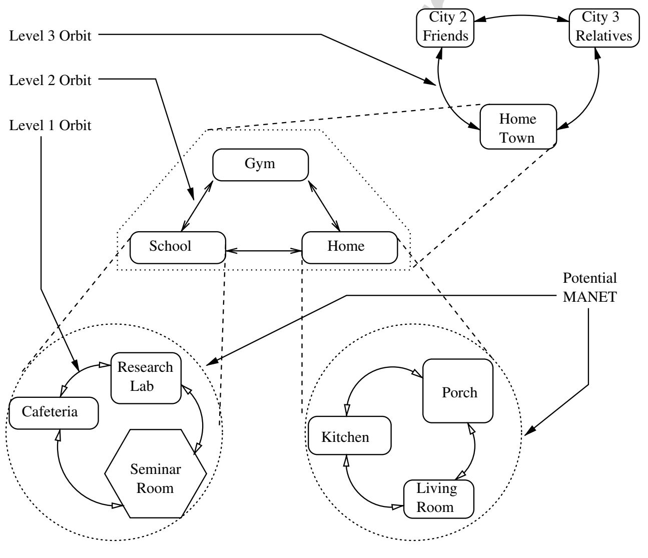

Fig. 1. A hierarchical view of sociological orbits.

2.1. Our ORBIT framework

In this section, we describe a simple yet practical macro-level mobility framework, and refer to it simply as ORBIT. As mentioned in Section 2, sociological movement pattern associates an individual node's movement with several special regions, or hubs within a terrain that the individual node visits. The ORBIT framework allows for the creation of a certain number of hubs within the simulation terrain for all the nodes, as specified by the parameter *Total Hubs*. Each node can visit a subset of hubs assigned to it, thereby creating its orbit. The list of hubs a node visits is bounded by *Hub List Size*, and the time it spends in each hub is specified by *Hub Stay Time*. The actual hub list assigned to a node along with the corresponding *Hub Stay Times* define an *Orbit*. ORBIT also allows for an occasional change in the specific list of hubs assigned to a node in its Orbit by defining an *Orbit Timeout*, upon which a node is assigned a fresh list of hubs to visit.

Without loss of generality, each hub may be considered to be a rectangle, with the length on each side bounded by *Hub Size*. The mobility pattern of individual nodes shall comprise of two parts: movement inside a hub, and movement in between hubs. The point to be noted here is that for each of the two parts, any known practical mobility models may be chosen, as suggested in Table 1. For detailed reference to these models, the reader is referred to Section 5.

In the following discussion, the movement inside each hub, which shall also be referred to as the intra-hub movement (IHM), is assumed to follow the Random Waypoint mobility model, whose speed range is denoted by *Intra-Hub Speed*, and whose pause time is denoted by *Intra-Hub Pause*.

Table 1
Example mobility models for ORBIT framework

Inside Hub	*Between* Hubs	Ref.
LMM	GMM	[9]
Area zone	City area	[10]
METMOD	NATMOD	[11]
Manhattan	Freeway	[1]

Table 2
ORBIT parameters

Category	Parameter
General	Total Hubs Hub List Size (min, max) Hub Size (min, max) Hub Stay Time (min, max) Orbit Timeout (min, max)
Inter-Hub	Inter-Hub Speed (min, max)
Intra-Hub (Random Waypoint)	Intra-Hub Pause Intra-Hub Speed (min, max)

Note that a non-zero minimum for the hub speed should be chosen to overcome the decaying average speed problem associated with Random Waypoint, as suggested in [30]. For Orbit movement (i.e., in between hubs), one may use a point-to-point linear (P2P Linear) model. In this model, when a node wants to move from one hub to another, it randomly selects a point within the destination hub and moves towards it linearly from its current position with a velocity defined by the range *Inter-Hub Speed*. Table 2 summarizes all of the Orbit parameters mentioned. Note that, our Orbit framework does not simply integrate two common mobility models (e.g., Random Waypoint, and P2P Linear), but most importantly, it also introduces the practical orbital movement in between hubs.

2.2. Example orbital models

By varying the placement of the different hubs within the terrain, and the hub-visit pattern of each node, we can use the ORBIT framework to generate some practically useful mobility models that suit different realistic scenarios. Note that, the following models only form an exemplary subset of mobility models that are capable of being generated by our ORBIT framework, while other models may also be obtained by varying the mobility models integrated within the framework (i.e., Random Waypoint and P2P Linear).

Random Orbit: By randomly placing hubs within the terrain, assigning randomly selected list of hubs to each node, and allowing each node to move randomly between the hubs assigned to it, a random orbit is formed for each node (as illustrated in Fig. 2). This model is general enough to be realistic, and specific enough to be useful. More specifically, it is suitable for modeling wireless devices carried by mobile users (e.g., graduate students, office employees) around a university/corporate campus, or attending technical conventions. As users move around, their wireless devices may record the hubs visited most often, and share the hub-based orbital mobility profile with trusted "acquaintances" either automatically, or with the user's permission/assistance. Such mobility profile can then help improve routing as described in Section 3.

Uniform Orbit: This model is similar to Random Orbit, except for the setup of the hubs. Unlike the previous model, the entire terrain is divided into a grid of hubs, without any hub overlapping another, as seen in Fig. 3(a). Such a model may be used to simulate scenarios like a school building, which may be completely divided into a set of non-overlapping classrooms, cafeterias, libraries, laboratories, etc., such that each unit may be considered as a hub. Students move along their Orbit between different rooms, and may change their Orbits upon Orbit Timeout every few weeks or months.

Restricted Orbit: This model is similar to Uniform Orbit, except for the hub-visit pattern of each node. In this model, each Orbit consists of a single hub, thereby restricting the movement of each node within a particular hub as illustrated in Fig. 3(b). The only time an Inter hub movement occurs is when an Orbit Timeout occurs, and a new hub is

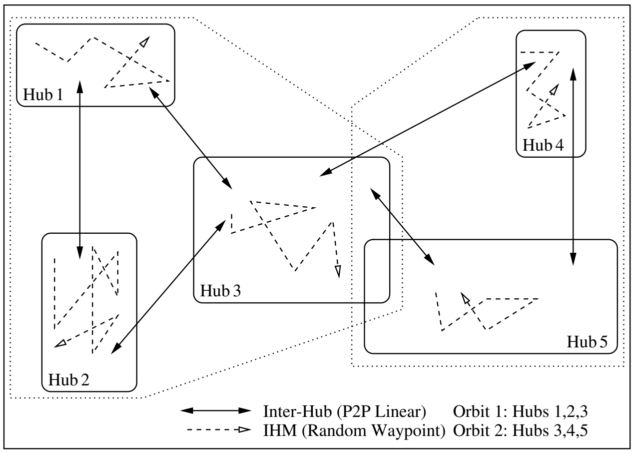

Fig. 2. The Random ORBIT model.

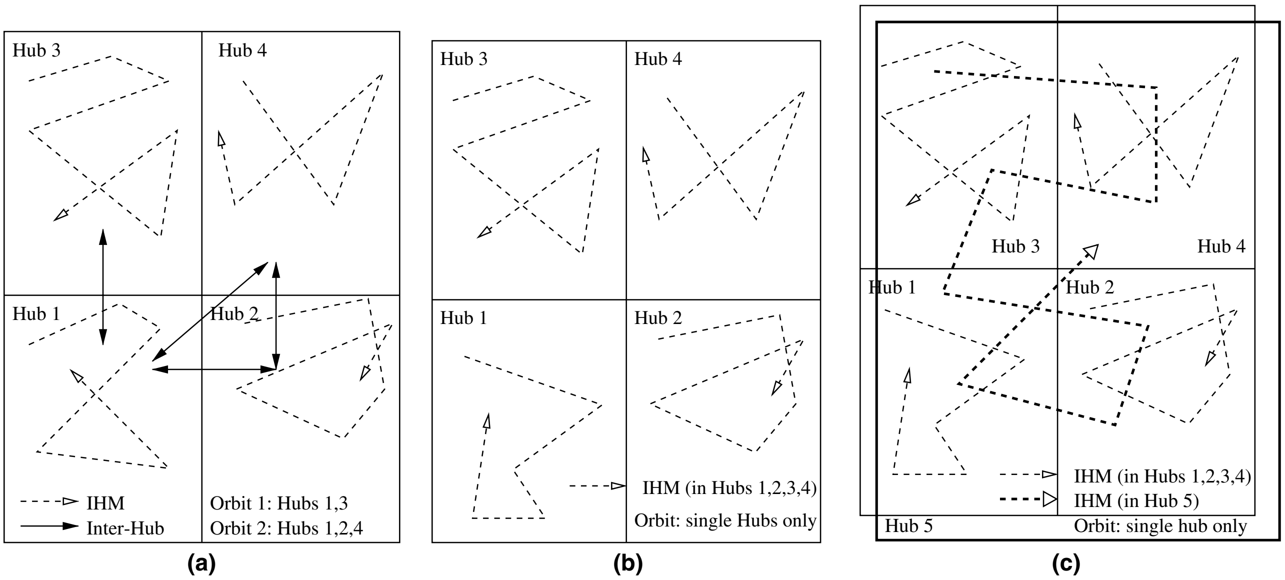

Fig. 3. Example orbital mobility models generated by ORBIT framework using Random Waypoint and P2P Linear. (a) Uniform orbit. (b) Restricted orbit. (c) Overlaid orbit.

assigned to the node. This model is useful in simulating office building scenarios, where not only is the entire terrain divided into several non-overlapping departments, but the movement of employees are also restricted within their respective departments. Over time, their work might require them to be reassigned to a different department, which can be modeled by the Orbit Timeout.

Overlaid Orbit: This model is identical to the Restricted Orbit, except for the addition of a special larger hub that is overlaid over all other smaller hubs, as illustrated in Fig. 3(c). Such a model may be useful when modeling an exhibition scenario where the entire terrain is divided into separate exhibition areas/booths, and nodes (presenters) in each booth are commonly restricted to their respective booth. In contrast, the mobility of attendees/viewers may span several such booths, thereby creating a larger overlaid hub. A more general variation is to have a larger hub that overlays several (but not all) smaller hubs that can be classified as belonging to some common group.

2.3. Analysis of orbital model characteristics

In this section, we analyze the models generated by the proposed ORBIT framework using a few of the protocol independent metrics defined in [1]. These metrics were shown to characterize the basic features of mobility models and their effect on protocol performance. Our motivation for this analysis is to illustrate that our example models provide multiple choices for modeling mobility under different scenarios described in the previous sub-section. We choose to vary the number of hubs in our study, since this also causes the hub sizes to vary simultaneously in all our grid based models except Random Orbit. We use a modified Random Waypoint model (which is modified to have a non-zero minimum as suggested in [30]) and the entire terrain (i.e., in a non-orbital movement) as a reference point in our simulations. We perform our simulations in GloMoSim [31] with 100 nodes in a $1000 \times 1000\ \mathrm{m}^2$ area for 1000 s. Each node is assumed to have a radio range of 250 m. Note that, for a protocol independent analysis of mobility models, no data traffic, routing protocol, or MAC layer needs to be considered. Values used for the ORBIT parameters are as follows:

- Total Hubs = Vary
- Hub List Size = 1 to Total Hubs
- Hub Size = 150–250 m (Random Orbit)
- Hub Stay Time = 50–100 s
- Orbit Timeout = 250–500 s
- Inter-Hub Speed = 10–30 m/s
- Intra-Hub Pause = 1 s
- Intra-Hub Speed = 1–10 m/s

To analyze our models, we use average degrees of spatial and temporal dependence as the mobility metrics, and the average number of link changes and link duration as the connectivity graph metrics. For a detailed description of the metrics, the reader is referred to [1]. No single metric serves as the determining factor in choosing one model over another.

2.3.1. Mobility metrics

The following metrics study the mobility characteristics of nodes within the same spatial locality, and that of each individual node within a small time interval.

Average degree of spatial dependence: Spatial dependence indicates the similarity in the velocities of two nodes that are within a specific range from each other, which is chosen to be $2 * (\textit{Radio Range})$ in our simulations. Since Restricted Orbit and Overlaid Orbit do not allow Orbit of nodes, nodes in the same hub follow the same hub parameters for a long time, resulting in a high spatial dependence as seen in Fig. 4(a). Random Orbit and Uniform Orbit support Orbit and hence have a lower value for this metric, while the modified Random Waypoint has an intermediate value.

Average degree of temporal dependence: Temporal dependence indicates the similarity in the velocities of a node within a specific time interval, which was taken to be 20 s in our simulations. In Random Orbit, the hub sizes are not affected by the number of hubs, unlike in the other models, but the amount of overlap among hubs increases sharply. Due to this overlapping, the path from one hub to another along an Orbit is short, causing quick changes in IHM leading to a low temporal dependence as seen in Fig. 4(b). In the modified Random Waypoint, nodes have a single IHM where a slower speed change results in a higher temporal dependency. For the remaining models, a larger number of hubs leads to a steady decrease in the temporal dependence with the number of hubs.

2.3.2. Connectivity graph metrics

The following metrics study the effect of mobility on the graph formed by the nodes and the radio links among them.

Average number of link changes: This is the average number of times a link between two nodes (that ever existed during the entire simulation) comes up from being down. Since in Restricted Orbit and Overlaid Orbit, nodes are confined to particular hubs, the probability that a link between two nodes in the same hub breaks (and then comes up later) is small. In the modified Random Waypoint, links that would break when two nodes move away have a low probability of coming up later as nodes move slowly in the entire terrain. In Random Orbit and Uniform Orbit, a link that is formed when two nodes visit a common hub may break when one of them moves away, but has a high probability of coming up again when they meet later in the common hub. In Random Orbit, these nodes may re-establish the link even if they are in different but overlapping hubs, thus showing the highest number of link changes in Fig. 5(a).

Average link duration: This is the average amount of time a link between two nodes stays up. The restricting nature of Restricted Orbit and Overlaid Orbit proves beneficial to link stability. Moreover, with an increase in the number of hubs, the hub sizes decrease causing the nodes to huddle even closer, increasing link duration as seen in Fig. 5(b). The Orbit supported by Random Orbit and Uniform Orbit causes link breaks to happen more often,

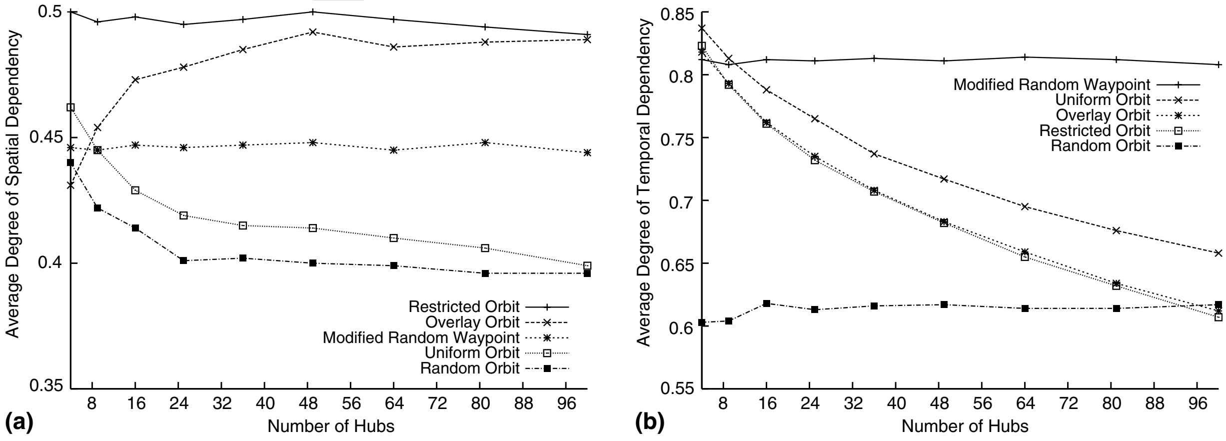

Fig. 4. Performance of mobility metrics vs. number of hubs. (a) Spatial, (b) Temporal dependence.

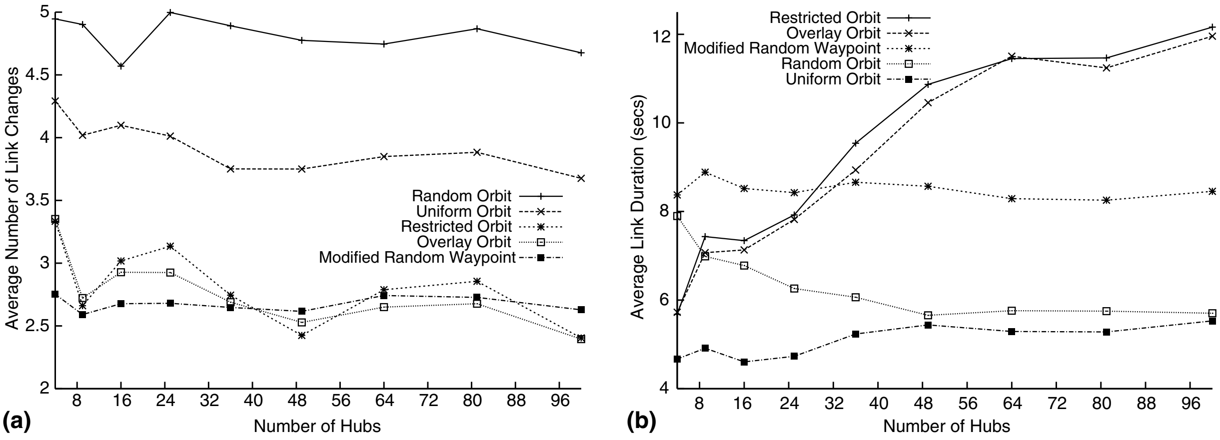

Fig. 5. Performance of connectivity graph metrics vs. number of hubs. (a) Link changes, (b) Link duration.

while the modified Random Waypoint shows an intermediate value.

The above study confirms that by integrating a macro-level orbital movement and other conventional micro-level mobility models (e.g., Random Waypoint, P2P Linear) our ORBIT framework is able to generate practical mobility models, that differ in terms of protocol independent metrics. Thus, such models can not only suit different practical scenarios, but also potentially affect routing protocol performances differently. For the remainder of this paper, we shall focus on the Random Orbit model when studying routing protocol performance.

3. Sociological orbit aware location approximation and routing (SOLAR)

In this section, we describe our sociological orbit aware location approximation and routing (SOLAR) protocol, which to the best of our knowledge, is among the first to make use of macro-level sociological mobility profiles of MANET users in obtaining approximate location information of mobile users as well as in improving routing. In the following discussion, nodes in SOLAR are assumed to be equipped with either GPS, or other localization devices[1] to facilitate greedy geographic forwarding as to be explained in detail in the following sections.

3.1. Protocol overview

Several motivations for *peer collaboration* (among acquaintances) were discussed by the authors in [32]. In one of our earlier work [33], we proposed an acquaintance based soft location management (ABSoLoM) scheme, and also showed the advantages of using acquaintances to form a distributed location database. Although one of the basic concepts in SOLAR is also the use of acquaintances, it differs from ABSoLoM in at least two significant ways. First, nodes in ABSoLoM use a formal acquaintanceship request and response among a few selected nodes that ensure the acquaintanceship to be mutual. Second, acquaintances not only cached each other's exact location coordinates, but also kept each other informed of their current coordinates through frequent location updates. In SOLAR however, acquaintanceship need not be mutual. As soon as one node gets to know of another node's hub list, it will treat that other node as an acquaintance and will cache its hub list information. This knowledge may be gained either by exchanging hub list information with that node directly when they are within radio range of each other, or through a trusted third party (e.g., common acquaintance).

Since the orbital mobility profile or, the hub list information stays valid for a much longer time when compared to the exact location coordinates, SOLAR can significantly reduce overhead in terms of *location updates* in the face of node mobility. More specifically, in SOLAR, each node only needs to know the terrain in terms of the hubs (i.e., their coordinates) in addition to its own location. Each node periodically sends its own coordinates and hub list in a *Hello* message to its immediate neighbors within radio range to facilitate both neighbor discovery (required for Inter-Hub geographic routing of packets) and hub list sharing (for forming

---

[1] Note that this assumption is common to most location aided MANET routing protocols suggested in literature.

new acquaintances). Only the acquaintances with an active data connection in between them need to notify each other by "location updates" when there is a change in any of their hub lists (as a result of an occasional Orbit Timeout). A more detailed description of SOLAR is as follows.

3.2. Information query propagation and response

In SOLAR, when a source has *data* to send and the destination is a neighbor within radio range, the *data* is directly transmitted. However, if the destination is not a neighbor, but is an acquaintance whose hub list is cached by the source, the *data* packet is sent towards the (center point of) hub(s) in that hub list (see Section 3.3 for more details). This packet transmission is based on "greedy geographic forwarding" [16], where each intermediate node chooses its next hop from amongst its neighbors who is closest to the destination than itself (see Section 3.4 for more details).

If no information about the destination's hub list is available, the source first selects a subset of its acquaintances in a way to be described in Section 3.5. For each chosen acquaintance, a separate *query* is sent to the hubs in that acquaintance's hub list. Such a transmission from a node to its acquaintance will be referred to as a *logical hop* here after, which often consists of multiple physical hops. An acquaintance responds to this *query* packet if it knows of a valid hub list of the destination. As an optimization, intermediate nodes (that are not acquaintances of the source) are also allowed to snoop into *query* packets and respond to them if possible.

If the acquaintance does not know the destination's hub list, it may forward the *query* to a subset of its own acquaintances, chosen appropriately as before. However, if the number of logical hops exceeds a specified threshold, the *query* packet is dropped by the acquaintance. If all the *query* packets are similarly dropped, the source will time out and may either drop the *data* packet, or retry sending new *query* packets to a different subset of its acquaintances, or resort to simple flooding of *query* packets.

If the *query* reaches the destination itself, it not only responds (with a *response* packet) with its own hub list, but also indicates its current hub. The current hub information is cached by the source and used for subsequent delivery of data in the same session. The cache timeout value for this current hub information will be based on the average Hub Stay Time of a node. Similarly, the hub list information itself will be cached at the source for a time proportional to the average Orbit Timeout.

3.3. Packet transmission to a hub list

Once the source of any packet (*query*, *response*, or *data*, as well as *update* as to be described later) knows the hub list information for that packet's destination, it first checks to see if the current hub information for that destination is available. If that information is unavailable, one copy of the packet is geographically forwarded towards (the center of) each of the hubs (i.e., simulcast) in the hub list of the destination. However, if the information is available, a single packet is geographically forwarded to (the center of) that current hub. In either case, the source inserts its own hub list and current hub information into the packet header to help the destination of that packet respond back to the source.

As mentioned earlier, location is updated only within nodes that have an active "data session" between themselves. More specifically, for *data* packets, when the first *data* packet is sent, a data session is considered active, which expires when no data is generated/sent within a specified interval. Throughout that active data session, the source keeps inserting its current hub and hub list information into each *data* packet to keep its location information updated at the destination. The destination of that active data session reciprocates with its current hub and hub list information in an *update* packet (which can double as an ACK) on getting the first *data* packet. From then on, whenever the destination moves out of its current hub, or starts to orbit a different hub list (on an Orbit Timeout), it notifies the source of the change by sending a *update* packet towards the current hub of the source. Since such *update* packets are restricted between the two ends of an active session only, they are sent out infrequently and incur little overhead.

Note that sending to the current hub is just an optimization attempt. If the destination is not in that current hub when the *data* arrives, the *data* can be cached by nodes in that hub for a limited amount of time. This will allow the destination to retrieve it later, if it visits that hub as part of its orbital movement. Just before the cached data is to be purged, the node that is closest to the center of the hub may simulcast copies of that data to the other hubs in the list of the destination. Of course, the source may also time out and decide to take an appropriate action (as discussed above).

3.4. Use of geographic forwarding for packet delivery

When the source of any packet wishes to send that packet to a hub (possibly containing the destination of that packet), it uses greedy geographic forwarding as mentioned before. As each intermediate node performs greedy geographic forwarding to push the packet closer to the intended hub's center coordinates, if there is no neighbor closer to the hub than the node itself (also called a local maxima or "geographic hole"), this node broadcasts the packet to all its neighbors. A neighbor in turn checks if any of its neighbors is closer to the hub than the intermediate node which started this broadcast. If it finds any such neighbor, it forwards the packet to it. Otherwise, it may either drop the data packet, or employ techniques to route around geographic holes as suggested in [17].

If the *Hub Size* is fairly large compared to the *Radio Transmission Range* of the nodes, once any packet reaches (i.e., enters) a hub before reaching the destination, an *Intra-Hub flooding* (of the packet by the nodes within the hub) is performed (as the exact coordinates of the destination may not be available). This is also done if the source itself lies in one of the hubs in the destination's hub list, in which case the source itself initiates the Intra-Hub flooding in an attempt to reach the destination. Such Intra-Hub flooding is not required when the Hub Size is comparable (or smaller) to the radio range since a packet can be overheard by all the nodes in the hub as it is geographically forwarded to the center of the hub.

In addition, an Intra-Hub flooding (if required) will introduce marginal overhead since a packet will only require to be flooded across a couple of radio hops to effectively cover the entire hub. To support such limited flooding, all packets are uniquely identified by a tuple (*source*, *destination*, *sequence id*), which enables nodes to identify and drop duplicate packets.

3.5. Querying a subset of acquaintances

A node may make a lot of acquaintances over its life time. Hence, to reduce the overhead due to the *query/response* packets, it needs to minimize the number of acquaintances it will query at any given time. On the other hand, since each acquaintance A_i covers (i.e., visits) a list of hubs H_i, this minimum subset of acquaintances need to be carefully chosen to maximize the coverage of hubs, thereby increasing the chances of obtaining the destination's information.

Let the hub list of an acquaintance A_i be denoted by $H_i = \{h_1, h_2, \ldots, h_m\}$, where each h_i is a particular hub. Let H be the set of hub lists $\{H_1, H_2, \ldots, H_n\}$ covered by a node's acquaintances $A_1, A_2, \ldots, A_n$. Let C be the set of hubs covered by all its acquaintances. That is, $C = H_1 \cup H_2 \cup \cdots \cup H_n$. Our problem is to find a minimum subset, $H' \subseteq H$ s.t.:

$$\forall h_i \in C,\ \exists H_j \in H',\ \text{s.t. } h_i \in H_j.$$

This is a minimum Set Cover problem, which is known to be NP-Complete [34]. To find an heuristic solution, we have adopted the Quine–McCluskey technique [35,36] used widely in Boolean Algebra for minimization of Boolean expressions. To describe this method, we first define a few terms as follows.

Prime acquaintance: An acquaintance A_i with hub list H_i is a *Prime* acquaintance if there is no other single acquaintance A_j whose hub list H_j includes H_i (i.e., $\nexists A_j$, s.t. $H_j \supseteqq H_i$). Formally, A_i (with H_i) is a *Prime* acquaintance *iff*:

$$\nexists A_j(\text{with } H_j), \quad \text{s.t. } \forall h_k \in H_i, \quad h_k \in H_i \Rightarrow h_k \in H_j.$$

Let us consider an example (1) where, $H_1 = \{1,2\}$, $H_2 = \{2,3,4\}$, $H_3 = \{1,4\}$, and $H_4 = \{3,4\}$, be the hub lists of acquaintances A_1, A_2, A_3, and A_4. Since none of A_2, A_3 or A_4 alone covers all the hubs of A_1, A_1 is a *Prime* acquaintance. Following the same principle, both A_2 and A_3 are also *Prime* acquaintances, whereas A_4 is not (since $H_2 \supseteqq H_4$).

Essential Prime acquaintance: This is a *Prime* acquaintance that covers at least one hub that is not covered by any other *Prime* acquaintance. Let $P = \{H_{p_1}, H_{p_2}, \ldots\}$ be the set of all the hub lists of *Prime* acquaintances $\{A_{p_1}, A_{p_2}, \ldots\}$. Then, a *Prime* acquaintance A_{p_i} with hub list H_{p_i} would be an *Essential Prime* acquaintance *iff*:

$$\exists h_k \in H_{p_i},\ \text{s.t. } \forall H_{p_j} \in P\ (j \neq i),\ h_k \notin H_{p_j}.$$

Continuing with the previous example 1, even though A_1 is a *Prime* acquaintance, it does not cover any hub that is not already covered by either A_2 or A_3. So A_1 is not an *Essential Prime* acquaintance. Following the same principle, A_3 is not an *Essential Prime* acquaintance either. However, A_2 covers hub

3 that is not covered by any other *Prime* acquaintance (i.e., A_1 and A_3). Although, A_4 covered hub 3, A_4 is *not* a *Prime* acquaintance, and hence ignored. Thus, A_2 is the only *Essential Prime* acquaintance in our example.

To query the optimal subset of acquaintances, a node first examines the hub lists of its acquaintances and determines its *Prime* and *Essential Prime* acquaintances. All the *Essential Prime* acquaintances are then chosen, and all the hubs in C that they cover are marked. If any hub in C is left unmarked, a non-essential *Prime* acquaintance covering the maximum number of unmarked hubs is chosen next, and the corresponding hubs are marked. This procedure is repeated by adding one more non-essential *Prime* acquaintance at a time, until all the hubs in C get marked.

Once the optimal subset of acquaintances is determined, each acquaintance in that subset is queried as explained in detail in Section 3.2. Referring to the previous example 1, first A_2 (being an *Essential Prime* acquaintance) will be chosen, following which any one of the other *Prime* acquaintances (A_1 or A_3) will be chosen to cover hub 1 (which is not covered by A_2). Moreover, to minimize the number of responses generated for a particular query, the source may "anycast" (send to any one of a list of destinations) query packets to hubs that are common to the list of multiple acquaintances. Thus, in our example if eventually A_1 and A_2 get selected, separate query packets will be sent to hub 1 for A_1 and to hubs 3 and 4 for A_2, but a single anycast packet destined for any of A_1 or A_2 will be sent to their common hub 2. In addition to reducing responses, this will also minimize the number of query packets generated, leading to a lower control overhead.

3.6. An example scenario

Continuing our previous example involving mobile users attending a large technical symposium, let us assume that three graduate students are at the same conference. The mobile wireless devices carried by them along with those carried by other convention attendees form the MANET shown in Fig. 6, where the Random Orbit model may be assumed for the sake of our discussion here. The different rooms shown in the figure are assumed to hold different conference tracks, a poster session, and an exhibition area that are held concurrently. In addition, there is a registration area, a lounge and a cafeteria. Suppose Student 1 frequents the rooms hosting the *Conference Track* 1, *Conference Track* 3, and *Posters* (which constitute Student 1's Orbit), while Student 2 frequents the *Lounge* and the room for *Conference Track* 4 (which constitute Student 2's Orbit), and Student 3 frequents the rooms for *Exhibits* and *Conference Track* 4 (which constitutes Student 3's Orbit). When Student 1 is in the *Posters* section and Student 2 is in the *Lounge* (note that these 2 Hubs/Rooms overlap), they came within each other's radio range and shared their own hub lists. Later, say Students 2 and 3 meet at

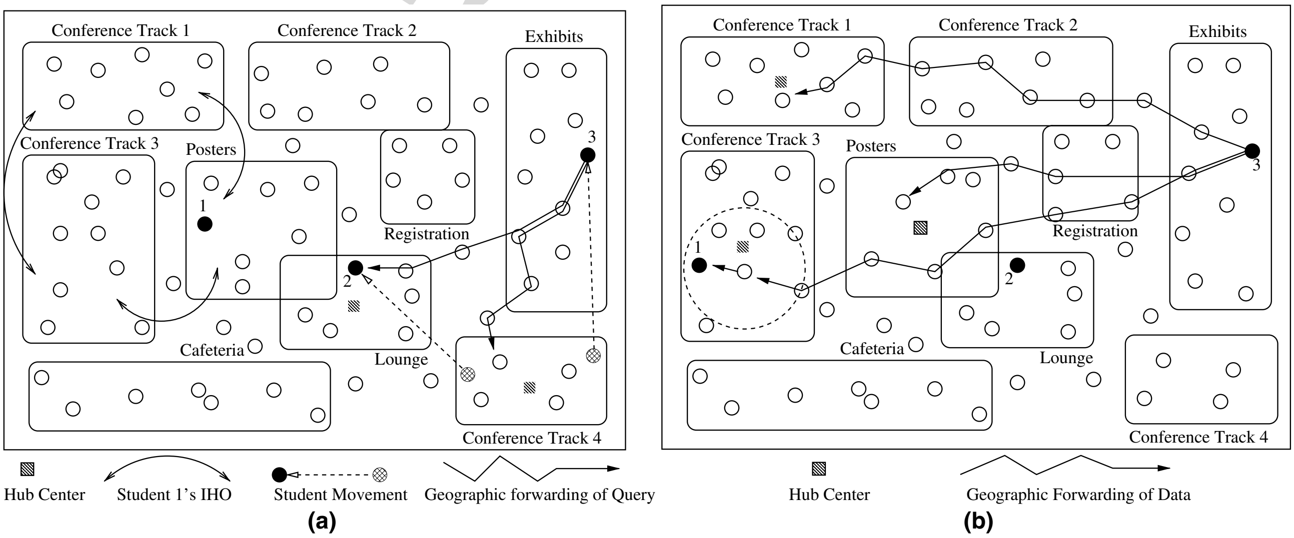

Fig. 6. Use of sociological mobility profiles of MANET users for location approximation and routing. (a) Student 3 queries acquaintance (Student 2) for Student 1. (b) Student 3 sends data to Student 1's hub list.

Conference Track 4 and also share their own hub lists. Later, if Student 3 wishes to locate Student 1 (whom he/she has not met yet), he/she can query his/her acquaintance (Student 2) for information related to Student 1's possible locations (as shown in Fig. 6(a)). Once Student 3 learns of Student 1's hub list, he/she can then simulcast messages geographically towards the hub list of Student 1 (as seen in Fig. 6(b)).

4. Performance analysis

In this section, to show that existing routing protocols under comparable assumptions (such as DSR and LAR) cannot be effectively used to deal with such practical orbital mobility pattern, we run extensive simulations to compare the performance of the SOLAR protocol with that of DSR and LAR scheme 1 (LAR1) using GloMoSim [31]. *Note that it is neither possible, nor our intention to compare our protocol with every other existing MANET routing protocols. Rather, the purpose of this exercise is to show that new routing protocols such as SOLAR is needed to take advantage of the orbital mobility pattern for efficient routing within MANET.* Accordingly, other routing protocols such as [37,17] that require the use of a separate location service above the routing layer are not considered here.

We implement two versions of the SOLAR protocol, SOLAR-1 and SOLAR-2. In SOLAR-1, a node sends *Hello* packets containing its own hub list to its 1-hop neighbors (i.e., nodes within its radio range), and only caches the hub lists of those neighbors. In SOLAR-2, each *Hello* packet also contains the hub lists of the 1-hop neighbors in addition to the node's own hub list. This allows nodes to cache the hub lists of the nodes that are either one or two radio hops away. In both versions of SOLAR, we use a threshold value of two for the number of *logical hops* any *query* packet may take before it is dropped. In this way, the query packets will only be sent to source's acquaintances, and their acquaintances. For comparison, we borrow the DSR and LAR1 implementations already available in the GloMoSim distribution.

For the simulation scenario, we consider a MANET built within a corporate campus consisting of several buildings (hubs). Corporate employees spend most of their time within the hubs and intermittently move in between hubs. To model realistic speeds of mobile users within such a MANET, we considered the work in [38–40]. We summarize real life speed values for various activities in Table 3 and fix the ORBIT Inter-Hub and Intra-Hub speed parameters accordingly, along with the other major simulation parameters as shown in Table 4. We chose three metrics to evaluate the performance of each protocol as described below.

Data throughput: This metric is defined as the ratio of the total number of data packets received correctly by all destinations, to the total number of data packets generated by all sources.

Relative control overhead: This metric is defined as the amount of control information (measured in bytes) that each node sends for each successfully received data packet in the network. For both LAR and DSR, we consider the *Route Request*, *Route Reply*, and *Route Error* packets as the control packets. In SOLAR, the control packets are *Hello*, *Hub List Query*, *Hub List Response*, and *Location Update* packets. Although, in SOLAR, the control packets

Table 3
Real life speed

Category	Type	Range (m/s)
Walking	Average	=1.34
	Olympic record	≈4.02
Running	Average	=4.00
	Olympic record	≈10.00
Cycling	Average	=8.94
	Olympic record	≈13.89

Table 4
Simulation parameters

General parameters	
Simulation duration (each run)	1000 s
Terrain size	1000m × 1000 m
Number of nodes (*users*)	*Vary* (default = 100)
Radio range	*Vary* (default = 250 m)
MAC protocol	IEEE 802.11
Mobility model	Random Orbit (RWP + P2P)
Orbit parameters	
Total Hubs (rooms)	*Vary* (default = 15)
Hub Size	*Vary* (default = 200–300 m)
Hub Stay Time	50–100 s
Orbit Timeout	250–500 s
Hub List Size	2 to Total Hubs
Inter-Hub Speed	*Vary* (default = 10–30 m/s)
Intra-Hub Pause	1 s
Intra-Hub Speed	1–10 m/s
Traffic parameters	
CBR connections	200 (five packets each) Random
Data payload	512 bytes per packet

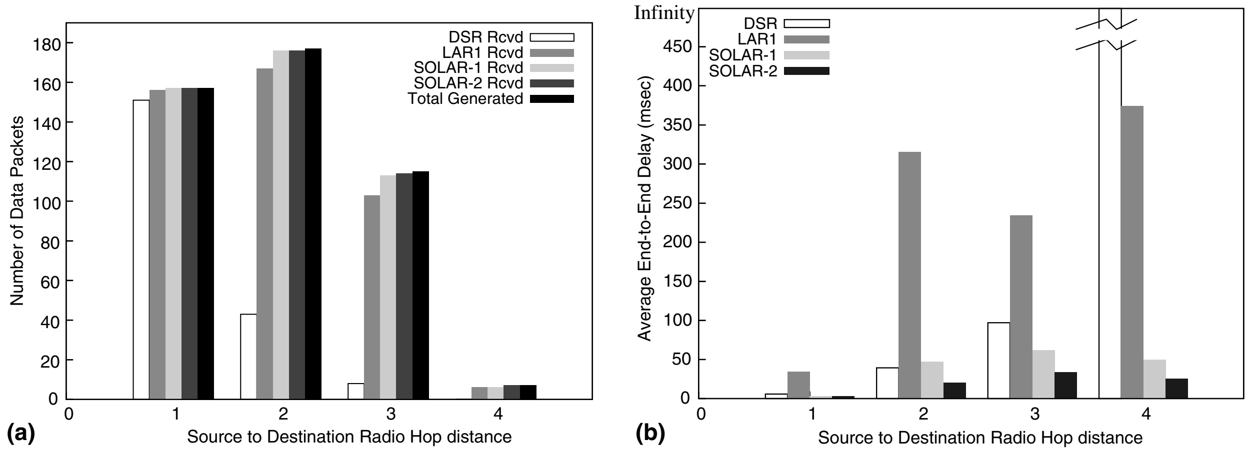

Fig. 7. Protocol performance at various source–destination radio hop distances. (a) Data delivered. (b) End-to-end delay (ms).

have larger size (in bytes) due to the mobility information contained in them, we show via simulations that both the overhead and delay are lower than those in DSR and LAR.

Approximation factor for end-to-end delay: The end-to-end delay measures the time from when a data packet is generated at the source, to the time when it is correctly received by the destination. Thus, this delay includes the discovery delay, which is the time taken to discover a path to the destination (in DSR and LAR), or the time taken to discover the destination's hub list (in SOLAR). Packets not delivered by any protocol are excluded from the calculation for that protocol, which may raise "fairness" concerns as to be discussed later. To compare different protocols as fairly as possible, we calculate the ratio of the delay observed in simulation and the minimum possible delay for a data packet in an ideal case, and call it the *approximation factor for end-to-end delay*. The minimum possible delay is the time taken by a packet, right after being generated, to make its way to the destination via minimum number of radio hops without any MAC contention, network queuing delay, etc. We use the same minimum possible delay while calculating the approximation factor for each of the three protocols.

The main reason for dividing the observed delay of each packet by the minimum possible delay of that packet is to account for the fact that different protocols can successfully deliver different data packets, and thus may incur different end-to-end delays. To further illustrate this point, we have run a few simulations with the default parameters mentioned in Table 4 and classified the data packets delivered successfully, in terms of the radio hop distances between their respective sources and the destinations. Fig. 7(a) shows the results of our simulation for a given set of packets generated with different source–destination radio hop distances (note that the number of packets generated with a distance of four hops is very small in this case). The number of data packets delivered by DSR drop significantly once the destination is more than two radio hops away from the source (whereas LAR1, SOLAR-1 and SOLAR-2 deliver a high percentage of packets consistently across all hops). On the other hand, the end-to-end delay in DSR increases with the hop distance (and may be considered to reach ∞ when the hop distance is 4 as zero packets are delivered in this case) as shown in Fig. 7(b).[2] Thus, a simple comparison of the average end-to-end delay, excluding these undelivered packets in particular protocols, would not be fair. More specifically, DSR would enjoy a lower average delay when compared to LAR1 just because DSR is unable to send the 4-hop packets which would otherwise incur a higher delay over a longer path. To introduce a sense of fairness in the delay performance comparisons, we choose to compare the *approximation factor*, which measures the end-to-end delay relative to the "optimum" delay (after the notion of "approximation factor" in algorithms that measure the closeness of an heuristic solution to the optimum solution).

In what follows, we will examine how different parameters such as total number of hubs (given a fixed terrain), Hub Size, Inter-Hub Speed, radio

---

[2] Whereas in LAR1 and SOLAR, this may not be the case due to the fact that the discovery delay (which is a part of the end-to-end delay) does not necessarily increase with the hop distance.

range, and the total number of nodes affect the protocol performance. To that end, we vary one of these five factors while fixing all others parameters.

4.1. Variation in total number of hubs

The number of hubs in the terrain affects protocol performance due to its direct impact on the expected node density within hubs, and the hub list sizes of each node, thereby affecting the protocol performance as described below.

Data throughput: Fig. 8(a) shows the data throughput of all the protocols with varying number of hubs. SOLAR-2 and SOLAR-1 perform the best with LAR1 showing comparable results. DSR has the lowest values for this metric.

The number of hubs seems to have little impact on SOLAR-2, SOLAR-1 and LAR1 but has an interesting impact on DSR. With a very few hubs, the number of nodes that happen to stay within each hub at any given time can be very large. This elevates the *broadcast storm problem* [41] (increased MAC layer contention) in DSR when flooding of discovery packets is attempted by any node, leading to unsuccessful route discovery and poor throughput. The performance of DSR improves with the number of hubs, but after a point, it deteriorates once again. This is because the hub list sizes of nodes also increases with the number of hubs, and as a result, the nodes enjoy greater freedom of movement within the terrain, adversely affecting DSR by increasing the chances of route failures.

LAR1 employs the caching of velocity and location information that helps in limiting the amount of flooding required, thereby resulting in much better performance. In the SOLAR protocols, as long as there is Inter-Hub movement whence the hub list information is shared amongst nodes, there is sufficient means to locate nodes and route packets to them, irrespective of the number of hubs.

Relative control overhead: From Fig. 8(b), we note that LAR1 has the highest overhead, followed by SOLAR-1, DSR and SOLAR-2, respectively. The majority of the overhead in flooding based protocols such as LAR1 and DSR is due to the route discovery process. Specifically, in LAR1, routes are discovered iteratively by increasing the size of the region where a destination is expected to be found. When the number of hubs is very low, they may be located far apart, requiring nodes to travel long distances as part of their Orbit. This leads to nodes moving out of LAR1's estimated region, causing repeated flooding and consequently increases the control overhead. On the other hand, if the number of hubs (and the hub list size along with it) is very large, nodes enjoy greater freedom of movement within the terrain. This too is not favorable for LAR1 for a similar reason as above. This is why a moderate number of hubs seems to result in a lower control overhead in LAR1.

DSR adopts a less aggressive flooding scheme and is shown to have a lower overhead than LAR1. In the case of SOLAR protocols, the periodic HELLO beacon in SOLAR-2 contains more information than that of SOLAR-1. Thus, the overhead in SOLAR-2 is more than that of SOLAR-1. More specifically, in SOLAR, hub lists stay valid for a longer time (than location coordinates, or routes), minimizing the number of query/response packets. In addition, the location update packets are also limited and infrequent. Thus, SOLAR protocols are able to maintain the lowest overhead among its competitors.

Approximation factor for end-to-end delay: The reasons given above also explain the approximation factor for delay of all the protocols as seen in

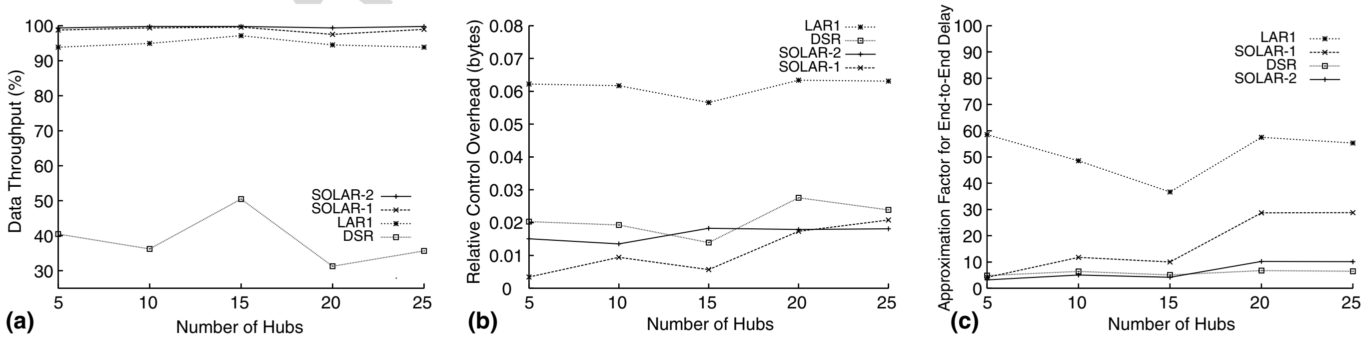

Fig. 8. Protocol performance vs. total number of hubs. (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

Fig. 8(c). LAR1 has the highest delay due to its iterative estimation of node location, and increased control overhead. In SOLAR-1, as the hub list size grows with the number of hubs, it takes a longer time to get the hub list of a destination with the assistance of only 1-hop neighbor information. Thus, the delay in SOLAR-1 increases marginally with increasing number of hubs. SOLAR-2, with more information, does considerably better than SOLAR-1, and is comparable to DSR. However, a point to note is that this delay in DSR is only averaged over the data packets it successfully received, which is far less than any other protocol. Overall, all protocols seem to perform the best with a moderate number of hubs for the default simulation terrain, hub size, and number of nodes. Accordingly, we set the default value of the number of hubs to 15 (see Table 4).

Note that the results on the relative performances of the three protocols shown in Fig. 8 are generally applicable to all other four cases to be described below where the hub size is fixed but one of the other four parameters varies, although the explanations may be slightly different in those four cases.

4.2. Variation in hub size

We study the effects of the hub size on the protocol performance in this section. In the following simulations, the hubs were considered to be square regions with equal sizes.

Data throughput: Fig. 9(a) shows that SOLAR-2, SOLAR-1, and LAR1 perform consistently well across all hub sizes, with the SOLAR protocols doing the best. On the other hand, since small hubs force nodes to stay very close to one another within a hub, DSR is adversely affected by the *broadcast storm* problem mentioned before, and hence does not perform well with small hub sizes. On the other hand, hub size has minimal effect on the throughput performance of SOLAR and LAR.

Relative control overhead: Once again, LAR has the highest control overhead, followed by DSR, SOLAR-2 and SOLAR-1, as seen in Fig. 9(b). The reasons are similar to those given in Section 4.1 for Fig. 8(b).

Approximation factor for end-to-end delay: Fig. 9(c) shows LAR1 to have the highest approximation factor for the end-to-end delay, with DSR and SOLAR-2 at a comparable minimum. When the hubs are small, there is hardly any overlap amongst them. Thus, if a node moves out of a hub, it most likely has to move a relatively long distance before reaching another hub. This has a negative impact on the location estimation of LAR1. On other hand, when the hubs are larger, there is a greater chance for hubs to overlap. Thus, even if a node moves to a new hub, its locality with respect to the terrain remains the same. This aids LAR1 in estimating node locations more accurately, and leads to a lower discovery delay with increasing hub size.

DSR's delay can also be negatively affected by MAC layer contention with a very small hub sizes, while SOLAR protocols enjoy a low hub list discovery latency as before, due to the use of the distributed location database within a network of acquaintances.

4.3. Variation in Inter-Hub Speed

By varying the *Inter-Hub Speed* we varied the amount of time nodes spend on average transiting in between hubs, with respect to their average Hub Stay Time. For the default *Inter-Hub Speed* range given in Table 4, the ratio of the Inter-Hub

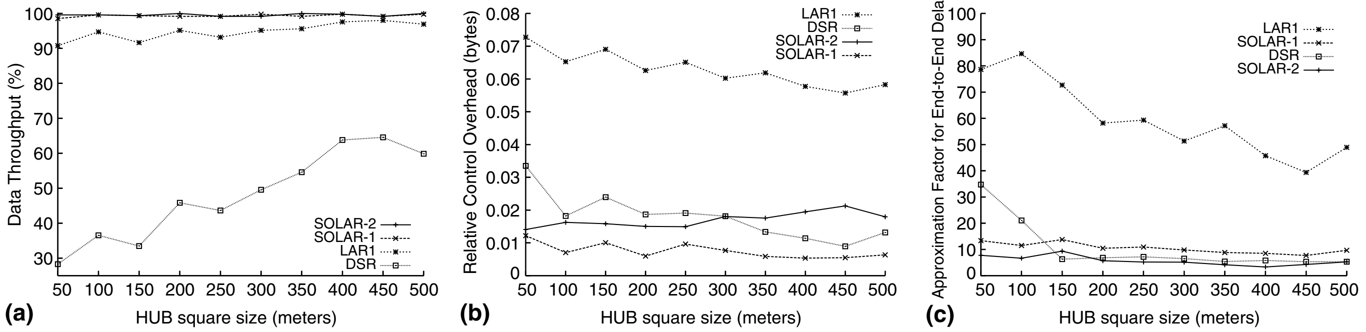

Fig. 9. Protocol performance vs. Hub Size. (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

Transit Time to Hub Stay Time is in between 0.3 and 0.15. We varied this speed from 2 m/s to 30 m/s so as to the change the value of this ratio from 2 to 0.15.

Data throughput: Fig. 10(a), shows that SOLAR-2, SOLAR-1, and LAR1 do consistently well for the entire range, while DSR performance fluctuates at several points. LAR1 manages to maintain a high throughput only at the cost of higher overhead and higher delay as confirmed by our earlier observations. In the SOLAR protocols, high values of the ratio (i.e., nodes spending a large amount of time traveling in between hubs) does not have a significant impact on the throughput performance. This is because in SOLAR, intermediate nodes also respond to queries, and cache data packets at each hub in the destination's hub list, in addition to being able to reach the destination outside a hub during geographic forwarding. On the other hand, lower values of the ratio (i.e., nodes spend considerable amount of time within hubs) only substantiates the practicality of the hub list information. DSR seems to be doing relative well for three different cases. First, when nodes spend more time outside hubs than inside, they have a low Inter-Hub Speed, in addition to the default low Intra-Hub Speed. This overall reduction in node velocity increases route stability in DSR, leading to good throughput. Second, when nodes spend most of their time within hubs, they move with low (default) Intra-Hub Speed leading once again to increased route stability. Third, DSR also seems to be doing well when nodes spend an equal amount of time inside and outside hubs, which leads to a more uniform node distribution that ultimately increases the chances of route discovery via flooding.

Relative control overhead: The relative performances of LAR1, SOLAR-2, and SOLAR-1 in Fig. 10(b) are similar to that observed in Fig. 9(b), and for similar reasons.

Approximation factor for end-to-end delay: In terms of this metric, LAR1 has the highest values while both SOLAR-2 and DSR show comparable minimum results in Fig. 10(c). DSR however, achieves this average approximation factor over a much smaller set of successfully delivered data packets when compared to the other three protocols. SOLAR-2 does better than SOLAR-1 as expected due to a higher amount of hub list caching. As node speed increases however, and they stay in hubs more often and travel very quickly in between hubs (i.e., lower ratio values), LAR may cache the lower Intra-Hub Speed and estimate a region around the last known hub. Thus, anytime a node moves out of a hub, LAR1 may fail to estimate the location correctly, thereby incurring higher delay with decreasing values of the ratio of Inter-Hub time to Hub Stay Time. This is also supported by marginal increase in LAR1 overhead, and marginal decrease in LAR1 throughput for smaller ratio values.

4.4. Variation in number of nodes (and data connections)

In this section, we study the effect of network load on our routing protocols by varying the number of nodes while keeping the number of connections per user constant, resulting in a varying total number of connections.

Data throughput: With a small number of nodes (and connections), LAR1 performs the best as shown in Fig. 11(a). In this case, DSR also benefits considerably and in fact, performs as well as SOLAR protocols. As for the SOLAR protocols, a very small number of nodes increases the chances

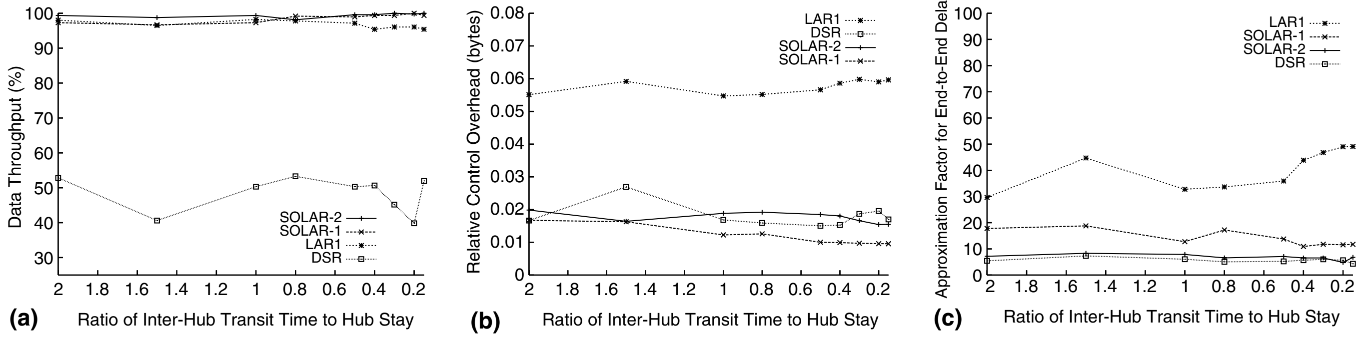

Fig. 10. Protocol performance vs. ratio of inter-hub transit time to hub stay time. (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

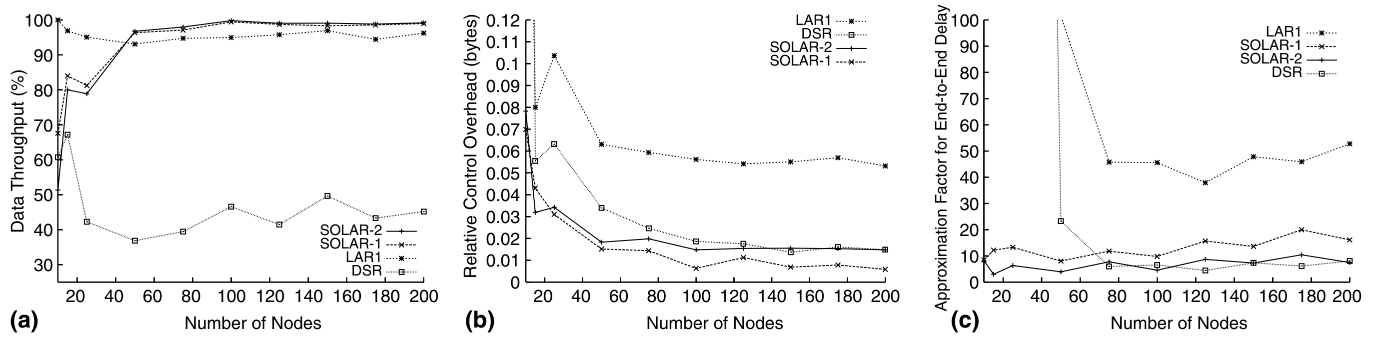

Fig. 11. Protocol performance vs. number of nodes (data connections = 2 × number of nodes). (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

of encountering a local maxima (or routing hole) while performing geographic forwarding. Additionally, with a fewer connections, SOLAR protocols can no longer benefit much by allowing nodes to learn other node's hub lists as they forward data packets for other nodes. Nonetheless, as the number of nodes (and data connections) increases beyond 40, SOLAR achieves highest throughput while DSR begins to get increasingly affected by the *broadcast storm* problem as discussed earlier.

Relative control overhead: As shown in Fig. 11(b), for all the protocols, the relative overhead reduces with increased number of nodes as they can make better use of the different information (path, location, velocity, hub list) cached in the intermediate nodes. The relative performance of the three protocols remains unchanged.

Approximation factor for end-to-end delay: Fig. 11(c) shows that both LAR1 and DSR have a significantly higher delays with a small number of nodes. This is because flooding becomes ineffective when there is a only small number of nodes that are restricted to move and stay within fixed hubs. On the other hand, in SOLAR protocols, the orbital mobility information of the nodes is still effective enough to keep the node discovery latency to a consistently low value.

4.5. Variation in radio range (and Hub Size)

The effect of a fixed radio range on varying hub sizes has already been discussed in Section 4.2. As a final study, in this section we scale the terrain up by varying the hub size and the radio range simultaneously, while retaining the default number of nodes and data connections.

Data throughput: As seen in Fig. 12(a), all protocols perform poorly with a smaller radio range. This can be explained as follows. In general, the average path length (in radio hops) between a source–destination pair increases with a smaller radio range. For LAR1 and DSR, the main impact of this effect is to increase the probability of a link failure, and ultimately leading to route failures. In the SOLAR protocols, a reduced radio range implies a lower number of radio neighbors who can continue greedy forwarding. This in turn increases the probability of failure due to the occurrence of local maxima in greedy forwarding. With larger radio ranges, all protocols perform much better as expected and

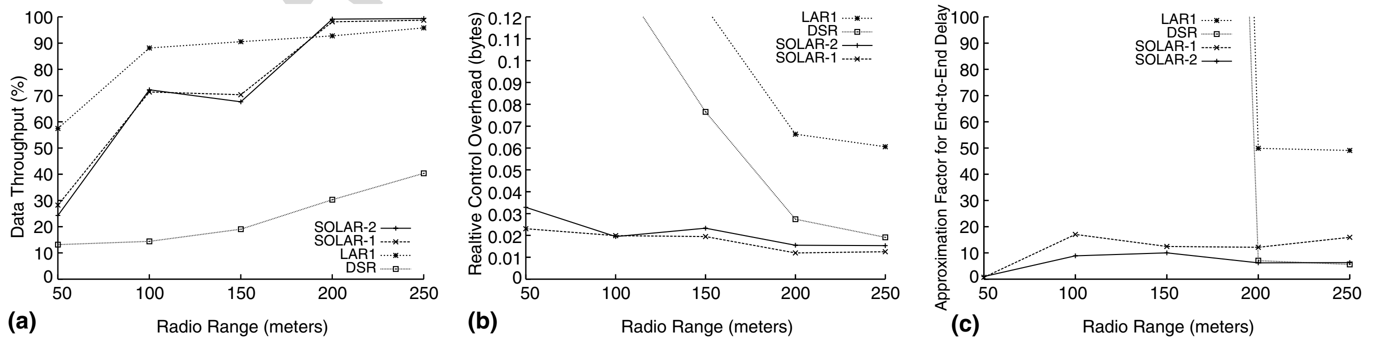

Fig. 12. Protocol performance vs. radio range (Hub Size = radio range). (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

the relative performances of these three protocols are consistent with the findings so far.

Relative control overhead: Due to an increased average path length caused by a shorter radio range, flooding based protocols will incur higher overhead and delays. This is confirmed in Fig. 12(b), which shows that both LAR1 and DSR have significant amount of control overhead for smaller radio range values. However, in the SOLAR protocols this effect is not as significant as in DSR or LAR1. More specifically, with shorter radio ranges, nodes in SOLAR have fewer neighbors that implies fewer acquaintances and a lower protocol maintenance overhead on average. On the other hand, for longer radio ranges, nodes get to know of many other nodes' hub lists, leading to reduced overhead in terms of request/response packets. Thus, the overhead in SOLAR is seen to be consistent across varying radio ranges.

Approximation factor for end-to-end delay: Fig. 12(c) shows that the delays for both LAR1 and DSR are significantly higher for small radio range values. This is because longer paths in source routing schemes tend to break more often, causing retransmissions, and in turn higher delays. In the SOLAR protocols, since nodes move along an orbit, they are able to continue to share hub lists with other nodes, and thus the delay remains consistently lower than that in DSR and LAR1.

An interesting point worth mentioning here is the relationship between the route/hub discovery latency in each protocol with the distance (in radio hops) from the source to the destination. In both LAR1 and DSR, the discovery latency for a previously unknown node is directly proportional to the distance. This is because, even with the caches at intermediate nodes, longer routes have a higher probability of link breaks, resulting in route errors and leading to higher discovery delay. However, in SOLAR protocols, this relation is not that intuitive. For example, in SOLAR-1, a source node may need to learn about a destination two hops away by querying an acquaintance that is say four hops away, thereby increasing the approximation factor for the end-to-end delay (i.e., the observed delay with respect to the ideal delay based on the distance between the source and the destination). On the other hand, it is equally likely that a source node may learn about a destination that is four hops away by simply querying a 1-hop neighbor that happens to be an acquaintance of the destination. More specifically, since the hub lists stay valid much longer than route caches, longer source-destination distances may still have end-to-end delays close to the ideal case due to reduced discovery latency. Thus, in SOLAR the discovery latency is tightly coupled with the knowledge of each node about other nodes' hub lists, and not as much dependent on the radio hop distance between the source and the destination as in LAR1 and DSR.

4.6. Effect of lossy links and gray zones

Wireless networks are seldom attributed with 100% link reliability due to varying channel conditions. As such, it is common to find asymmetric links and "gray zones" [42] within the network. Unlike a geographic hole, a gray zone may occur due to the presence of a radio neighbor during neighbor discovery, but failure during the actual transmission of "unicast" data to that neighbor.

Although the presence of specific gray zones affect only a fraction of links in the network, we decided to study the effect of uniform link loss probability across the network. Thus, the effect of link loss is quite significant even at very low loss probabilities as any link is subject to the probabilistic failure at any time.

Fig. 13(a) shows a steady decline in data throughput of all the protocols, with SOLAR protocols doing much better than DSR and LAR1. LAR1 seems to be affected the most. It is possible that due to the orbital movement of nodes, the cached information of LAR1 expires quickly, and causes nodes to restart the query process. As all links have a probability of failure, queries may be lost in gray zones more often leading to reduced throughput. On the other hand, SOLAR enjoys longer lifetime of cached hub list information and require less amount of query packets to be forwarded. This is also reflected in Fig. 13(b), where LAR1 is seen to have the highest relative overhead, and SOLAR the least.

The approximation factor for end-to-end delay in SOLAR is seen to increase in Fig. 13(c), while that in DSR and LAR1 is seen to decrease. This may be explained by the fact that nodes in SOLAR are able to deliver more packets than any of the other protocols, at the cost of additional delay incurred when data is cached at different nodes.

Thus, even though SOLAR is affected by increasing link loss probabilities, it still outperforms DSR and LAR1 in terms of data throughput and relative control overhead, while the increase in the

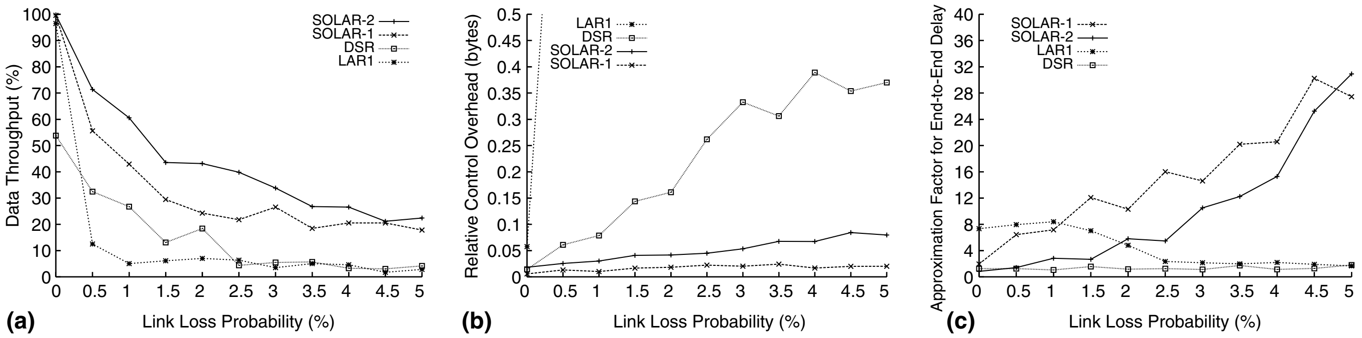

Fig. 13. Protocol performance vs. link loss probability. (a) Data throughput (%). (b) Relative control overhead. (c) Approximation factor for E2E delay.

approximation factor for delay is justifiable. Note that the techniques suggested in [42] may be adopted within SOLAR to reduce the occurrence of gray zones.

To summarize, based on the above study, we can firmly claim that while DSR and LAR make trade-offs between throughput, control overhead and delay, SOLAR is far superior to any one of these protocols in terms of higher data throughput, lower control overhead, and shorter end-to-end delay.

5. Comparison with related work

As discussed earlier in Section 1, the growing awareness of the effect of node mobility on protocol performance has led to research on practical mobility models and mobility adaptive routing schemes. A popular mobility model which is used exhaustively in mobile network research, Random Waypoint, is clearly not very realistic and has practical limitations as pointed out in [30]. In [43], the authors suggested enhancements to the model itself, including the use of acceleration to smoothen changes in speed and direction. While the authors in [3] focused on the application of voronoi graphs to model mobility in face of obstacles, those in [4], integrated three sub-models: perception, behavioral and movement, to simulate the mobility of each individual node as a close interaction of simple behavioral traits. To guarantee a steady state in node movement distributions, the authors in [5] used *renewal theory*, while authors in [44] introduced the concept of *stochastic correlation* in their VUM (variable user mobility) model for cellular systems.

Although, the previously described models could simulate the mobility of individual nodes, they could not capture the motion of entire groups in scenarios such as military drills, disaster recovery, search party etc. Accordingly, the authors in [6,7] proposed a mobility model called Reference Point Group Mobility, where an existing group leader determines a group's collective movement, while other members move independently within a small speed and angle deviation from that of the leader. Their work was later extended to include a mobility vector model, which smoothened the sudden changes in speed and direction. In [8], the authors surveyed several such *Entity based* (e.g., Boundless area, Gauss–Markov) and *Group based* (e.g., Column, Nomadic, Pursue) mobility models for ad hoc networks. In [1], the authors formalized a framework for analyzing mobility models in terms of protocol independent metrics, and also proposed the *Manhattan* and *Freeway* models to suit city-wide traffic.

Literature also describes several proposals that model mobility as a hierarchy. In [9], the authors suggested two hierarchical layers for a wireless ATM network: a deterministic global mobility model (GMM) to describe inter-cell movements, and a stochastic local mobility model (LMM) to describe intra-cell movements. In [10], the authors applied *transportation theory* to model: *City Area*, *Area Zone*, and *Street Unit*, at three hierarchical levels. On the same note, the authors in [11] proposed the Metropolitan (*METMOD*), National (*NATMOD*) and International (*INTMOD*) mobility models to model movements within metropolitan areas, in between them, and in between countries respectively. Despite this wide range of study on practical mobility, no prior work has studied the mobility of MANET users in between certain geographic regions of some social significance (which we call hubs).

Meanwhile, work has also been done on routing protocols to counter the adverse effects of node mobility. For instance, flooding based source routing schemes [2,45] employ aggressive route caching techniques to reduce the route discovery delay.

The proposed protocol in [12] also cached the node velocity to restrict the flooding of node discovery packets to a smaller geographic region containing the destination node. Observing that node locations could similarly aid in routing, several location management schemes [16,46,18,19] coupled with geographic routing were proposed for stateless routing in ad hoc networks. However, these schemes require accurate knowledge of locations (via the use of a GPS receiver, for instance), and the distributed management of accurate locations result in frequent location updates and protocol maintenance packets.

While the researchers in [47,48] studied the effects of various realistic scenarios on the existing MANET routing protocols, the authors in [13,14] suggested methods like a connected virtual backbone to help routing protocols adapt to node mobility. Moving one step further, researchers in [25,26,24] started to focus on mobility pattern/information awareness in routing protocols. Their method was to use such mobility pattern information obtained either via continuous location tracking, or micro-level mobility prediction to help make low level routing decisions such as the best choice of next hop to take. On the other hand, our work focusses on the macro-level mobility information for location approximation and routing at the hub level.

The authors in [29] acknowledged the higher probability of a node visiting a location which it has frequented the most in the past. They developed a *delivery predictability* metric to perform probabilistic routing for intermittently connected networks like the delay tolerant networks (DTN). However, unlike us, they did not attempt to profile the user mobility based on a collection of regions frequented by mobile users in some periodic (but partially deterministic) sequence. To the best of our knowledge, we are the first to both model such a sociological mobility framework for a MANET user, and exploit the hub level mobility pattern in making location approximation and routing decisions.

6. Conclusion

In this work, we have exploited a partially deterministic and hub-based orbital mobility information by observing the social influence on the macro-mobility pattern of each MANET user. Specifically, we have proposed an ORBIT framework to capture the orbital movement pattern for MANET users based on a list of places or hubs that they frequently visit. ORBIT is useful in capturing realistic mobility by integrating micro-level mobility models with a macro-level inter-hub orbit. We have also used this simple yet practical mobility information to perform intelligent routing. In particular, we have proposed a sociological orbit aware location approximation and routing (SOLAR) protocol for MANET and established the advantages of SOLAR over conventional MANET routing protocols such as LAR and DSR in terms of higher data throughput, lower control overhead, and lower end-to-end delay.

References

[1] F. Bai, N. Sadagopan, A. Helmy, Important: a framework to systematically analyze the impact of mobility on performance of routing protocols for ad hoc networks, Proceedings of IEEE INFOCOM 2 (2003) 825–835.

[2] J. Broch, D.B. Johnson, D.A. Maltz, Dynamic Source Routing in Ad hoc Wireless Networks, Mobile Computing, Kluwer Academic Publishers, 1996, pp. 153–181.

[3] A. Jardosh, E.M. Belding-Royer, K.C. Almeroth, S. Suri, Towards realistic mobility models for mobile ad hoc networks, Proceedings of ACM/IEEE MobiCom (2003) 217–229.

[4] D. Tan, S. Zhou, J. Ho, J. Mehta, H. Tanabe, Design and evaluation of an individually simulated mobility model in wireless ad hoc networks, in: Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDSMSC), San Antonio, TX, 2002.

[5] G. Lin, G. Noubir, R. Rajaraman, Mobility models for ad hoc network simulation, Proceedings of IEEE INFOCOM (2004).

[6] X. Hong, M. Gerla, G. Pei, C.-C. Chiang, A group mobility model for ad hoc wireless networks, Proceedings of ACM/IEEE MSWIM (1999) 53–60.

[7] X. Hong, T. Kwon, M. Gerla, D. Gu, G. Pei, A mobility framework for ad hoc wireless networks, in: ACM 2nd International Conference on Mobile Data Management (MDM), 2001, pp. 185–196.

[8] T. Camp, J. Boleng, V. Davies, A survey of mobility models for ad hoc network research, wireless communications and mobile computing (WCMC): Special issue on mobile ad hoc networking: Research, Trends and Applications 2 (5) (2002) 483–502.

[9] T. Liu, P. Bahl, I. Chlamtac, Mobility modeling, location tracking, and trajectory prediction in wireless atm networks, IEEE Journal on Selected Areas in Communications 16 (6) (1998) 922–936.

[10] J. Markoulidakis, G. Lyberopoulos, D. Tsirkas, E. Sykas, Mobility modeling in third-generation mobile telecommunications systems, IEEE Personal Communications 4 (4) (1997) 41–56.

[11] D. Lam, D. Cox, J. Widom, Teletraffic modeling for personal communications services, IEEE Communications Magazine (1997) 79–87.

[12] Y.B. Ko, N.H. Vaidya, Location-aided routing (LAR) in mobile ad hoc networks, Proceedings of IEEE INFOCOM'98 (1998) 66–75.

[13] S. Basagni, D. Turgut, S. Das, Mobility-adaptive protocols for managing large ad hoc networks, in: Proceedings of IEEE International Conference on Communications (ICC), 2001, pp. 1539–1543.

[14] S. Butenko, X. Cheng, D.-Z. Du, P. Pardalos, On the construction of virtual backbone for ad hoc wireless networks, Cooperative Control: Models, Applications and Algorithms (2003) 43–54.

[15] P. Enge, P. Misra, Special issue on global positioning system, Proceedings of the IEEE 87 (1) (1999) 3–15.

[16] J. Li, J. Janotti, D.S.J.D. Couto, D.R. Karger, R. Morris, A scalable location service for geographic ad hoc routing, Proceedings of ACM/IEEE MobiCom'00 (2000) 120–130.

[17] B. Karp, H.T. Kung, GPSR: Greedy perimeter stateless routing for wireless networks, Proceedings of ACM/IEEE MobiCom'00 (2000) 243–254.

[18] C.T. Cheng, H.L. Lemberg, S.J. Philip, E. van den Berg, T. Zhang, SLALoM: a scalable location management scheme for large mobile ad-hoc networks, in: Proceedings of Wireless Communications and Networking Conference, vol. 2, 2002, pp. 574–578.

[19] S. Philip, C. Qiao, Hierarchical grid location management for large wireless ad hoc networks, ACM SIGMOBILE Mobile Computing and Communications Review (MC2R) 7 (3) (2003) 33–34.

[20] J. Li, C. Blake, D.D. Couto, H.I. Lee, R. Morris, Capacity of ad hoc wireless networks, in: Proceedings of the 7th ACM International Conference on Mobile Computing and Networking, vol. 10 (4), 2001, pp. 61–69.

[21] P. Gupta, P.R. Kumar, The capacity of wireless networks, IEEE Transactions on Information Theory 46 (2) (2002) 388–404.

[22] M. Gastpar, M. Vetterli, On the capacity of wireless networks: the relay case, IEEE INFOCOM'02 3 (2002) 1577–1586.

[23] M. Grossglauser, D.N.C. Tse, Mobility increases the capacity of ad hoc wireless networks, IEEE/ACM Transactions on Networking 10 (4) (2002) 477–486.

[24] S. Samal, Mobility pattern aware routing in mobile ad hoc networks, MS thesis, Virginia Polytechnic Institute and State University, May 2003.

[25] W. Su, S.-J. Lee, M. Gerla, Mobility prediction and routing in ad hoc wireless networks, International Journal of Network Management 11 (1) (2001) 3–30.

[26] W. Wang, I.F. Akyildiz, On the estimation of user mobility pattern for location tracking in wireless networks, Proceedings of IEEE Globecom'02 (2002) 619–623.

[27] S. Jain, K. Fall, R. Patra, Routing in a delay tolerant network, Proceedings of ACM SIGCOMM'04 34 (4) (2004) 145–158.

[28] A. Vahdat, D. Becker, Epidemic routing for partially connected ad hoc networks, Technical Report CS-2000-06, Duke University, April 2000.

[29] A. Lindgren, A. Doria, O. Schelen, Poster: Probabilistic routing in intermittently connected networks, in: Proceedings of The Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2003).

[30] J. Yoon, M. Liu, B. Noble, Random Waypoint considered harmful, Proceedings of IEEE INFOCOM'03 2 (2003) 1312–1321.

[31] X. Zeng, R. Bagrodia, M. Gerla, Glomosim: a library for parallel simulation of large-scale wireless networks, in: Proceedings of the 12th Workshop on Parallel and Distributed Simulations (PADS)'98, 1998, pp. 154–161.

[32] S. Banerjee, P. Chrysanthis, Peer support in a mobile world, in: Proceedings of NSF Workshop on Context-Aware Mobile Database Management, January 2002.

[33] J. Ghosh, S. Philip, C. Qiao, Acquaintance based soft location management (ABSoLoM) in MANET, in: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)'04, vol. 1, 2004, pp. 166–171.

[34] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of np-Completeness, W.H. Freeman and Company, 1979.

[35] W.V. Quine, A way to simplify truth functions, American Mathematical Monthly 62 (9) (1955) 627–631.

[36] E.J. McCluskey, Minimization of Boolean functions, Bell Systems Technical Journal 35 (5) (1956) 1417–1444.

[37] J. Li, P. Mohapatra, LAKER: Location aided knowledge extraction routing for mobile ad hoc networks, in: Proceedings of Wireless Communications and Networking Conference, vol. 2, 2003, pp. 1180–1184.

[38] R.L. Knoblauch, M.T. Pietrucha, M. Nitzburg, Field studies of pedestrian walking speed and start-up time, Transportation Research Board Records No. 1538, 1996.

[39] R.M. Queen, H.-Y. Lin, M.T. Gross, Standardized Running Speed versus Self-selected Running Speed: A Between-Trial and Between-Day Reliability Study, The Center for Human Movement Science—UNC Chapel Hill, October 2003.

[40] A. Williams, Cycling Speed, Article in Peak Performance, No. 65. Available from: <http://www.pponline.co.uk/encyc/0065.htm>.

[41] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, J.-P. Sheu, The broadcast storm problem in a mobile ad hoc network, in: Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking, 1999, pp. 151–162.

[42] H. Lundgren, E. Nordstrom, C. Tschudin, The gray zone problem in IEEE 802.11b based ad hoc networks, ACM SIGMOBILE Mobile Computing and Communications Review (MC2R) 6 (3) (2002) 104–105.

[43] C. Bettstetter, Smooth is better than sharp: a random mobility model for simulation of wireless networks, in: Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems, 2001, pp. 19–27.

[44] K. Lee, S. Kim, Modeling variable user mobility with stochastic correlation concept, Computer Networks: The International Journal of Computer and Telecommunications Networking 38 (5) (2002) 603–612.

[45] C.E. Perkins, E.M. Royer, Ad hoc on-demand distance vector routing, in: Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, 1999, pp. 90–100.

[46] S.-C. Woo, S. Singh, Scalable routing protocol for ad hoc networks, Wireless Networks 7 (2001) 513–529.

[47] G. Ravikiran, S. Singh, Influence of mobility models on the performance of routing protocols in ad-hoc wireless networks, in: Proceedings of IEEE Vehicular Technology Conference, vol. 4, 2004, pp. 2185–2189.

[48] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, M. Degermark, Scenario-based performance analysis of routing protocols for mobile ad-hoc networks, in: Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking, 1999, pp. 195–206.

Joy Ghosh completed his Bachelors in Computer Science and Engineering from Jadavpur University, Calcutta, India in 1999. He worked for Cadence Design Systems till 2001 in the Systems and Functional Verification team. Thereafter, he shifted his focus to Wireless Networking issues in the areas of Ad hoc, Sensor and Cellular networks. He received his M.S. in Computer Science from the State University of New York at Buffalo in 2003. Currently he is a Ph.D. candidate in the same department under Dr. Chunming Qiao with special interests in routing, location management, mobility profiling and analysis. He has worked as an intern in Fujitsu Laboratories being part of the SuperComm 2005 project. He has also worked as a summer intern in Nokia Research developing a performance engineering tool for GPRS networks.

Sumesh J. Philip received the B. Tech. degree in Electronics and Communication Engineering from the National Institute of Technology, Calicut, India, in 1997, the M.S. and Ph.D. degrees in Computer Science and Engineering from the University at Buffalo (SUNY), NY, in 2001 and 2005, respectively. He is currently an Assistant Professor in the Computer Science Department at Western Illinois University, Macomb. He has worked as a summer intern in Telcordia Technologies, Nokia Research and Bell Laboratories investigating issues related to wireless and cellular networks. His research interests include location and mobility management in mobile ad hoc and cellular networks, resource efficient routing in sensor networks, location tracking and its applications, and performance analysis of network protocols.

Chunming Qiao is a Professor at the University at Buffalo (SUNY), where he directs the Lab for Advanced Network Design, Analysis, and Research, (LANDER) that conducts cutting-edge research work on optical networks, wireless networks, and the Internet. He has published more than 120 papers in leading technical journals and conference proceedings, and is recognized for his pioneering research on Optical Internet, in particular, the optical burst switching (OBS) paradigm. His work on integrated cellular and ad hoc networking systems (iCAR) is also internationally acclaimed and has been featured in Business Weeks and Wireless Europe.

He is on the editorial board of several journals and magazines including IEEE Communications and IEEE/ACM Transactions on Networking (ToN), and have guest-edited three IEEE JSAC issues and one issue of ACM/Baltzer's Mobile Networks and Applications Journal (MONET). He has chaired and co-chaired many international conferences and workshops including the High-Speed Networking Workshop (formerly GBN) at Infocom'01 and Infocom'02, Opticomm'02, and the symposium on Optical Networks at ICC'03.