This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2007 proceedings.

Modeling Time-variant User Mobility in Wireless Mobile Networks

Wei-jen Hsu*, Thrasyvoulos Spyropoulos†, Konstantinos Psounis‡ and Ahmed Helmy*
*Dept. of Computer and Information Science and Engineering, University of Florida, Gainesville, Florida 32611-6120
†INRIA, Sophia-Antipolis, France
‡Dept. of Electrical Engineering, University of Southern California, Los Angeles, California 90089-2560
Email: wjhsu@ufl.edu, Thrasyvoulos.Spyropoulos@sophia.inria.fr, kpsounis@usc.edu, helmy@cise.ufl.edu

***Abstract*— Realistic mobility models are important to understand the performance of routing protocols in wireless ad hoc networks, especially when mobility-assisted routing schemes are employed, which is the case, for example, in delay-tolerant networks (DTNs). In mobility-assisted routing, messages are stored in mobile nodes and carried across the network with nodal mobility. Hence, the delay involved in message delivery is tightly coupled with the properties of nodal mobility.**

Currently, commonly used mobility models are simplistic random *i.i.d.* model that do not reflect realistic mobility characteristics. In this paper we propose a novel *time-variant community mobility model*. In this model, we define communities that are visited often by the nodes to capture *skewed location visiting preferences*, and use time periods with different mobility parameters to create *periodical re-appearance* of nodes at the same location. We have clearly observed these two properties based on analysis of empirical WLAN traces. In addition to the proposal of a realistic mobility model, we derive analytical expressions to highlight the impact on the *hitting time* and *meeting times* if these mobility characteristics are incorporated. These quantities in turn determine the packet delivery delay in mobility-assisted routing settings. Simulation studies show our expressions have error always under 20%, and in 80% of studied cases under 10%.

I. Introduction

In recent years, there has been an exponential growth in the popularity of portable computation and communication devices. Advances in wireless communication technologies and standards have made ubiquitous communication an emerging reality. With the ever expanding deployment of these wireless-capable devices, there is an increasing interest in a new communication paradigm and applications that are made possible through the new opportunities.

Ad hoc networks are self-organized, infrastructure-less networks consist of only wireless devices. In traditional ad hoc networks, it is generally assumed, albeit implicitly, that communications between nodes occur through multi-hop, complete paths in space. However, this assumption is in question for several reasons. First, multi-hop spatial routing increases the number of transmissions and channel contention, and hence reduces the capacity of scarce wireless bandwidth [6]. Second, such end-to-end paths may not always exist, given the wide variation of potential adverse settings (e.g., low node density, unpredictable mobility) in which wireless communication may take place. Due to the fore-mentioned reasons, routing schemes falling under the general framework of *mobility-assisted routing* have been proposed recently, as a measure to improve the wireless network capacity[7] and increase the feasibility of communication in more challenged environments[1].

Mobility-assisted routing schemes, as opposed to path-based ad hoc routing protocols, utilize nodal mobility to disseminate messages in the network. In mobility-assisted routing, transmissions from the senders to the receivers are not always completed immediately through a connected, complete multi-hop path. Rather, when a sender moves to close proximity of some other nodes in the network, the packet is forwarded to and stored in these intermediate nodes for potentially long time periods, waiting for the transmission opportunities to other nodes in the network. Instead of being considered as a detrimental factor that makes reliable communication difficult, mobility provides communication opportunities in mobility-assisted routing. Hence, in these settings, mobility and nodal encounter are crucial components to understand the network performance.

However, most research studies on mobility-assisted routing assume simplistic mobility models, such as the random walk[3], [4], [5] (in general, *i.i.d.* models), or a priori knowledge of future mobility[2]. These assumptions provide scenarios amenable to mathematical analysis that provides good insights to system performance. However, these simple mobility models do not address the complexity of nodal mobility in real-life settings. In all these models, all mobile nodes behave statistically identical to each other, and their behaviors do not change with respect to time. As the underlying mobility model is an important factor of the performance of mobility-assisted routing schemes, there is an increasing need for mobility models that capture the realistic mobility characteristics *and* remain mathematically manageable.

Our main contribution in this paper is the proposal of a *time-variant community mobility model*. The model captures several important mobility characteristics we observed from empirical wireless LAN (WLAN) traces. Specifically, we utilize the WLAN traces from the archives at [22] and [23] to understand the prominent mobility characteristics of current wireless network users in university campuses and corporate buildings. We have identified *skewed location visiting preferences* and *periodical re-appearance at the same location* as two prominent trends existing in multiple traces[12]. These mobility characteristics are central in our daily activities but have not been addressed by existing mobility models. In the proposed model, we create communities to serve as popular locations for the nodes, and implement time periods in which the nodes move differently to induce periodical behavior. To our best knowledge, this is the first mobility model that captures non-homogeneous behavior in both *space* and *time*.

Moreover, the proposed *time-variant community model* can be mathematically treated to derive analytical expressions for two important quantities of interest that determine the performance of mobility-assisted routing schemes: the *hitting time* and the *meeting time*. Starting from stationary nodal distribution, the *hitting time* is the average time before a node moves towards the vicinity of a randomly chosen geographical location, and the *meeting time* is the average time before two nodes move to the vicinity of each other. These quantities capture the time between available communication opportunities under the mobility model, and can be used as building blocks to analyze the performance of more complex packet forwarding schemes [4], [5]. We further show that our theoretical derivation is accurate through simulation cases with a wide range of parameter sets. In all cases, the error between simulation results and theoretical values is less than 15% for the *hitting time* and 20% for the *meeting time*, and for 80% of studied cases the error is below 10%.

The remaining of the paper is organized as follows: In section II we discuss related work. For clarity, a simplified version of our *time-variant community mobility model* is introduced in section III, and the expressions for the *hitting time* and the *meeting time* are derived in section IV-A and IV-B, respectively, and validated with simulation in section IV-C. In section V, we show a good matching between the mobility characteristics of our model and the real WLAN traces and further discuss about the possible extensions of the *time-variant community model*. We conclude the paper in section VI.

II. Related Work

Mobility-assisted routing has been proposed for various purposes in the literature. In [7], a two-hop routing scheme has been shown to improve the network capacity. In [1], it is proposed to overcome the intermittent communication opportunity in challenged network settings (generally known as delay-tolerant networks (DTNs)), with low node density or unpredictable mobility. So far, most studies on mobility assisted-routing in the literature assume either a complete knowledge of future mobility and encounters[2] or an *i.i.d.* mobility pattern[7] with each individual node following simple mobility models, such as a random walk[3], [4], [5], mainly for the sake of theoretical tractability. In [17] the authors assume a constant meeting rate between the mobile nodes to derive the inter-meeting times. These previous works focus on deriving the performance of mobility-assisted routing with mathematical analysis. In this work we extend the scope of the analysis by proposing a *time-variant community model*, in which the nodal movement preferences are not *i.i.d.* in space and not homogeneous across time.

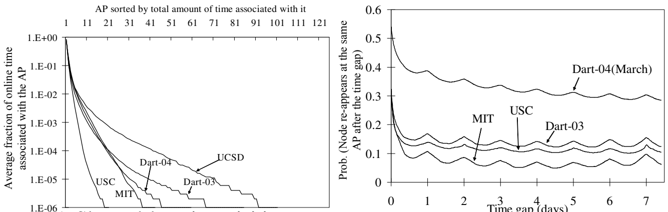
Along a different line of research, to understand mobility empirically, there has been WLAN measurement works which reveal the important mobility characteristics of the real-world wireless network users [9], [10], [11]. Large-scale deployments of WLANs in university and corporate campuses provide an excellent platform in which huge amount of user data can be collected and analyzed. Communities for WLAN trace-related study are available at [23] and [22].

We combine the two streams of research in this paper by taking into account the mobility properties observed in WLAN users and proposing a mathematical manageable mobility model. In [12] we identified several prominent properties that are common in multiple WLAN traces collected from various environments, including on-off behavior, skewed location visiting preferences, and periodical re-appearing behavior of nodes. Hence, we believe a good mobility model for wireless network users should preserve these characteristics. In this paper we extend the concept of community model proposed in our previous paper [16] further to include time-variant, non-*i.i.d.* behavior of mobile nodes.

There are several previous attempts to build models for WLAN users with the properties observed in WLAN traces[13], [14], [15]. These models match the preference and the pause duration with the users in observed traces. They fall into the category of the WLAN association model, in which currently associated access points are used as the indicator for mobile node locations. Hence the applicability of the models is specific to WLANs. However, in the case of mobility-assisted routing, nodes communicate not only when they are associated with fixed network infrastructure (such as WLAN), but also when they meet when moving between places. Hence we abstract the mobility characteristics from WLAN traces, and propose a continuous mobility model (i.e., node locations are given by (x, y) coordinates) which has wider applicability. Note that it would be more relevant to compare our model to such continuous human mobility traces. However, due to the unavailability of such traces, our best choice is WLAN-based ones. Moreover, modeling the exact trajectory may not be as crucial to our goal as providing a continuum in the mobility process.

There are several other efforts to collect mobile node encounter traces with hand-held devices [20], [21]. In these works the performances of routing protocols are directly obtained by finding the inter-encounter time distribution, instead of being derived from a mobility model. This approach is different but complementary to our approach. However, note that it is possible to derive encounter patterns from mobility models, but not verse visa, so we choose to focus on the more fundamental task (i.e., formulate a realistic mobility model). We are interested in comparing the inter-meeting times of our model with these traces in the future.

The concept of community is also mentioned in [19]. However, the authors assume the attraction of a community to a mobile node is derived from the number of friends of this



(a) Skewed location visiting preferences.

(b) Periodical re-appearance at the same location.

Fig. 1. Two important mobility features observed from WLAN traces.

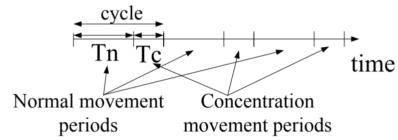


Fig. 2. Alternating normal movement periods and concentration movement periods in the time-variant community mobility model.

node currently residing in the community. In our paper we assume that the nodes follow location-based preference to make movement decisions, and each node moves independently of the others. Mobility models with inter-node dependency require a solid understanding for the social network structure, which is an important area under development. We choose to leave this as future work.

III. Description of Time-variant Mobility Model

A. Mobility Characteristics Observed in WLAN Traces

The main objective of the paper is to propose a mobility model that captures the important mobility characteristics observed from daily lives. To understand mobility realistically, we consult with the wireless LAN traces collected by several research groups (e.g., traces available at [22] or [23]). We acknowledge that these observations are made from WLAN traces, which do not register continuous movement of the devices. In particular, the devices are not always on, and typically there are some gaps in the coverage of access points (APs) in these networks. Nonetheless, we believe that two important features we observed from the traces[12], *skewed location visiting preferences* and *periodical re-appearance at the same location*, should remain in our proposed mobility model. This belief is based on the observations of typical daily activities of human beings: Most of us tend to spend most of the time at a handful of frequently visited locations, and a recurrent daily or weekly schedule is an inseparable part of our lives. In Fig. 1(a) we display *skewed location visiting preferences* property observed from WLAN traces, where a node on average spends more than 65% of its online time with one AP, and more than 95% of online time at as few as 5 APs. In Fig. 1(b) we display *periodical re-appearance at the same location* property, where a node re-appears at the same AP with higher probability after a time-gap of integer multiples of days. These are the two major features we wish to retain in the proposed *time-variant community mobility model*.

In order to capture *skewed location visiting preferences*, we need to define popular location(s) for the nodes in the continuous mobility model. We achieve so by defining the *communities* for each node and making a node visits its own community more often than other areas. Note that the notion of community is created to describe the most visited geographical area of a node (e.g., most visited building on campus), and different nodes can pick different communities and hence do not behave identical to one another (i.e., non-*i.i.d.* behavior in space). In order to capture *periodical re-appearance at the same location*, we establish structures in time, by defining time periods during which nodes go to their communities with even higher probability. Note that such structure in time not only creates periodicity, but also naturally captures the omnipresent time-dependent behavior in our daily lives, e.g., people go to offices during working hour, restaurants for lunch during noon time, and home after work with higher probability. This is along the same line suggested in [18]: Time-variant behavior has to be explicitly modeled in order to capture the changes of behavior across time.

B. Time-variant Community Mobility Model

In this section we describe a simplified version of the *time-variant community mobility model* which consists of two types of time periods and a single community in each time period, for the sake of clear presentation. In spite of its simplicity, it is sufficient to capture the major trends of the two fore-mentioned mobility features. The model can be extended to model more complex mobility, as we show in section V.

To implement time structure, in our model time is divided into two types of periods: *Normal movement periods (NMP)* of fixed lengths T_n, and *concentration movement periods (CMP)* of fixed lengths T_c. The later is created to capture the periodical re-appearance behavior shown in Fig. 1(b): We can assign high probability for a node to visit its community during the CMP, so it re-appears at its community with high probability with period $(T_n + T_c)$. These two types of time periods occur alternatively, as illustrated in Fig. 2. In the following discussion, we assume that the mobility model starts from the starting boundary of normal movement period at $t = 0$. Within each time period, we assign a community to each node. The community serves as an abstraction of the frequently visited location for the node during this period of time. The communities are much smaller than the N-by-N simulation area. We denote the community size in the NMP as a C_n-by-C_n square and the community size in the CMP as a C_c-by-C_c square. We assume that the community locations are chosen at random within the simulation area. Note that, the detailed parameters chosen for our model depend heavily on the targeted scenario, while we focus on proposing a *generic* model in the following derivations.

In each time period, a node has two different modes of movement: *local epoch* and *roaming epoch*. In a local epoch, the mobility of the node is confined within its community. In a roaming epoch, the node is free to move in the whole simulation area. At the beginning of each epoch, the node picks a speed uniformly distributed between $[v_{min}, v_{max}]$ and

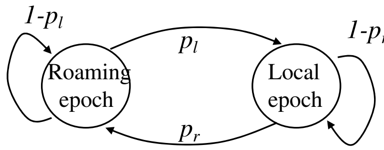


Fig. 3. Illustration of the 2-state Markov chain model of local/roaming epoch.

TABLE I

PARAMETERS OF THE TIME-VARIANT COMMUNITY MOBILITY MODEL

N	Edge length of simulation area
C	Edge length of community
$v_{min}, v_{max}, \overline{v}$	Minimum, maximum, and average of movement speed
$T_{max}, \overline{T}$	Maximum and average pause time after each epoch
$\overline{L}$	Average epoch length
$p_r(p_l)$	Probabilities of choosing a roaming (local) epoch after a local (roaming) epoch
$\pi_l(\pi_r)$	Stationary distribution of percentage of local (or roaming) epoch among all epochs
$P_{move}(P_{pause})$	Stationary distribution of fraction of time the node is in moving (or pause) status
K	Transmission range of nodes
HT_{case}	Expected hitting time under the given "case"
MT_{case}	Expected meeting time under the given "case"

a movement angle uniformly between $[0, 2\pi]$, and performs a constant speed, random direction movement in the corresponding area. The movement length of each epoch is drawn from an exponential distribution with average epoch length of $\overline{L}$. If the node hits the boundary of the simulation area in a roaming epoch, or hits the community boundary in a local epoch, it is re-inserted from the other end of the area (i.e., the boundaries are "torus" boundaries). At the end of the epoch, the node picks a pause time at the end point uniformly from $[0, T_{max}]$, and picks the movement mode for the next epoch according to a two-state Markov model, as shown in Fig. 3. For example, if the node has just finished a local epoch, it chooses the next epoch to be roaming with probability p_r or local with probability $1 - p_r$. These probabilities can be set differently in NMP and CMP, hence we denote them as p_{rn} and p_{rc}, respectively. Note that in the remainder of the paper we may append subscript l or r to indicate the parameters for *local* or *roaming* epochs and subscript n or c to indicate the parameters for *NMP* or *CMP*. We summarize the notations we used to describe the mobility model in Table I. Also note that these parameters can be set differently for different nodes to induce heterogeneous nodal behavior. However, for simplicity, we leave the details as future work.

Lemma 3.1: At any given time instant, the node is in one of the following four states: (a) moving in a local epoch, (b) moving in a roaming epoch, (c) pause after a local epoch, (d) pause after a roaming epoch. The stationary distribution of probability in each state is:

$$P_{move,l} = \frac{\pi_l(\overline{L_l}/\overline{v_l})}{\pi_l(\overline{L_l}/\overline{v_l} + \overline{T_l}) + \pi_r(\overline{L_r}/\overline{v_r} + \overline{T_r})}, \quad (1)$$

$$P_{move,r} = \frac{\pi_r(\overline{L_r}/\overline{v_r})}{\pi_l(\overline{L_l}/\overline{v_l} + \overline{T_l}) + \pi_r(\overline{L_r}/\overline{v_r} + \overline{T_r})}, \quad (2)$$

$$P_{pause,l} = \frac{\pi_l\overline{T_l}}{\pi_l(\overline{L_l}/\overline{v_l} + \overline{T_l}) + \pi_r(\overline{L_r}/\overline{v_r} + \overline{T_r})}, \quad (3)$$

$$P_{pause,r} = \frac{\pi_r\overline{T_r}}{\pi_l(\overline{L_l}/\overline{v_l} + \overline{T_l}) + \pi_r(\overline{L_r}/\overline{v_r} + \overline{T_r})}, \quad (4)$$

where $\pi_r = p_r/(p_r + p_l)$ and $\pi_l = p_l/(p_r + p_l)$. The above stationary probabilities can be calculated for NMP and CMP independently.

Proof: The derivation follows easily from calculating the stationary distribution of the local and the roaming epochs using basic Markov chain theory[8], and from taking into account the average duration of each epoch. ■

IV. DERIVATION OF THEORETIC EXPRESSIONS

In this section we derive the theoretic expressions of the hitting and the meeting times under the time-variant community model and validate it with simulation. The expected *hitting time* is the average time for a node, starting from the stationary distribution, to move into the transmission range of a fixed, randomly chosen target coordinate (i.e., "hits" the target) in the simulation area. The expected *meeting time* is the expected time for two mobile nodes, both starting from the stationary distribution, to move into the transmission range of each other. Note that this definition is different from *inter-meeting time* used in [17], [20], which accounts for the time duration between a meeting event and the subsequent one. Both quantities have important implications for the message delivery time in mobility-assisted routing scenarios. The hitting time is important when we have some fixed points in the network for the mobile nodes to collect information (e.g., moving sensors find an event of interest occurring at random location) or deliver information (e.g., drop of its readings to fixed sink). The meeting time is especially important if store-and-carry message forwarding across multiple mobile nodes is an application of interest. It is the average time before two given nodes encounter with each other, and hence it is closely related to the packet delivery delay in sparse networks where nodal encounters are the main opportunities for message delivery.

A. Derivation of the Hitting Time

The sketch of the derivation of the hitting time is as follows: We first derive the probability of hitting the target in a unit-time slice, P_h, under the mobility model. Since the node is performing random direction movements and the target is also chosen at random, each time unit can be considered as an independent Bernoulli trial with success probability P_h. We can then calculate the expected hitting time using the geometric distribution. The details are given below.

With the existence of communities, whether the chosen target is in one of the communities changes the expected hitting time value significantly. The hitting times for four possible cases are hence calculated separately, and then take weighted average. In the following, the first and the second subscripts in $HT_{i,j}$, where $i, j \in \{in, out\}$, correspond to whether the target is in the community during the NMP and the CMP, respectively. $P_{i,j}$ is the probability of the corresponding $HT_{i,j}$.

Lemma 4.1:

$$HT_{overall} = P_{in,in}HT_{in,in} + P_{in,out}HT_{in,out} + P_{out,in}HT_{out,in} + P_{out,out}HT_{out,out}, \tag{5}$$

where $P_{in,in} = (C_n^2/N^2)(C_c^2/N^2)$, $P_{in,out} = (C_n^2/N^2)(1 - C_c^2/N^2)$, $P_{out,in} = (1 - C_n^2/N^2)(C_c^2/N^2)$, *and* $P_{out,out} = (1 - C_n^2/N^2)(1 - C_c^2/N^2)$.

Proof: The proof follows from basic probability and simple geometric arguments. Note that communities are chosen independently at random for both the NMP and the CMP. ■

In order to calculate the expected hitting time for each case, the unit-time hitting probability should be calculated separately. For the sake of simplicity, we only show the details of deriving $HT_{out,in}$ in the following Lemma 4.2 and Theorem 4.4. The hitting time for the other three cases can be derived in a similar manner.

Lemma 4.2: The unit-time hitting probabilities for the case $\{out, in\}$ *in NMP and CMP are:*

$$P_{h,n} = P_{move,r,n}(2K\overline{v_r}/N^2) \quad and \tag{6}$$

$$P_{h,c} = P_{move,r,c}(2K\overline{v_r}/N^2) + P_{move,l,c}(2K\overline{v_{l,c}}/N_c^2). \tag{7}$$

Proof: The hitting event can only occur when the node is physically moving, as a non-moving node cannot encounter the fixed target which is not in its transmission range. Since the target location is chosen out of the community for the NMP, it can only be hit during roaming epochs in the NMP. We neglect the small probability that the target is chosen out of the community but close to it, so the node can actually hit the target during a local epoch. When a node moves with average speed $\overline{v_r}$, on average it covers a new area of $2K\overline{v_r}$ in unit time. Since random direction movement covers the whole simulation area with equal probability [16], and the target coordinate is chosen at random, it falls in this newly covered area with probability $2K\overline{v_r}/N^2$, and hence the unit-time hitting probability is $P_{h,n} = P_{move,r,n}(2K\overline{v_r}/N^2)$, i.e., when the node performs a roaming movement *and* the target is in the newly covered area in the time-unit.

The target location is chosen in the community for the CMP, so it can be hit during both roaming and local epochs. The hitting probability is the sum of these two cases. For roaming epochs the argument is the same as in the last paragraph, and for local epochs, the node covers randomly chosen new area of $2K\overline{v_{l,c}}$ in the community, hence it hits the target with probability $2K\overline{v_{l,c}}/N_c^2$.

■

Note that the movement made in each time unit does not increase or decrease the probability of hitting the target in subsequent time units, therefore each time unit can be considered as an independent Bernoulli trial with success probability given in Lemma 4.2. The corollary below immediately follows.

Corollary 4.3: The probability for at least one hitting event to occur in the whole NMP and CMP, respectively, are:

$$P_{H,n} = 1 - (1 - P_{h,n})^{T_n}, \quad P_{H,c} = 1 - (1 - P_{h,c})^{T_c}. \tag{8}$$

Theorem 4.4: The expected hitting time for the $\{out, in\}$ *case can be calculated as:*

$$\begin{aligned} HT_{out,in} &= Prob(hit\ in\ NMP)HT_{out,in,hit\ in\ NMP} \\ &+ Prob(hit\ in\ CMP)HT_{out,in,hit\ in\ CMP} \\ &= \frac{P_{H,n}}{P}(T_n + T_c)(1/P - 1) \\ &+ \frac{P_{H,n}}{P}\frac{1 - (1 + T_nP_{h,n})(1 - P_{h,n})^{T_n}}{P_{h,n}(1 - (1 - P_{h,n})^{T_n})} \\ &+ \frac{(1 - P_{H,n})P_{H,c}}{P}(\frac{T_n + T_c}{P} - T_c) \\ &+ \frac{(1 - P_{H,n})P_{H,c}}{P}\frac{1 - (1 + T_cP_{h,c})(1 - P_{h,c})^{T_c}}{P_{h,c}(1 - (1 - P_{h,c})^{T_c})}, \end{aligned} \tag{9}$$

where $P = P_{H,n} + P_{H,c} - P_{H,n}P_{H,c}$ *is the probability for at least one hitting event to occur in one full cycle of NMP and CMP.*

Proof: The expected hitting time for the $\{out, in\}$ case can be calculated as the weighted average of two separate subcases: The first hitting event occurs during NMP or CMP. We can view the occurrence of hitting events in two types of time periods as independent coin toss trials, which give head with probability $P_{H,n}$ and $P_{H,c}$, respectively. We want to calculate the number of flips needed until we get the first head, when we flip these two coins alternatively, starting from the NMP.

The success probability for each full cycle (containing one NMP and the subsequent CMP) is $P = P_{H,n} + P_{H,c} - P_{H,n}P_{H,c}$. The probabilities for the first hitting event to occur in NMP and CMP are $P_{H,n}/P$ and $(1 - P_{H,n})P_{H,c}/P$, respectively, as in each cycle of time periods NMP precedes CMP, and the time period structure is repetitive in itself.

For each sub-case, the time until the first hitting event can be further divided into two parts: The complete time periods before the last time period in which the hitting event occurs, and the fraction of the last time period until the hitting event occurs. If the hitting event occurs during the NMP, the expected duration of whole periods before that is $(T_n + T_c)(1/P - 1)$, since it takes $1/P$ full cycles for one success event to occur if the success probability is P. The time until the first hitting event occurs in the last NMP is

$$\begin{aligned} &\sum_{i=1}^{Tn} i\frac{P_{h,n}(1 - P_{h,n})^{i-1}}{1 - (1 - P_{h,n})^{T_n}} \\ &= \frac{1 - (1 + T_nP_{h,n})(1 - P_{h,n})^{T_n}}{P_{h,n}(1 - (1 - P_{h,n})^{T_n})} \approx \frac{1}{P_{h,n}}. \end{aligned} \tag{10}$$

The last approximation holds if T_n is large.

If the first hitting event occurs in the CMP, the expected duration of whole periods before that is $(T_n + T_c)(1/P - 1) + T_n$, and the remaining fraction of the last CMP can be calculated in a similar fashion to (10). Putting all these components together, we arrive at (9). ■

Finally, the overall expected hitting time is derived by solving all four cases in (5) following the procedure outlined in Lemma 4.2 and Theorem 4.4. Note that the last equation for the expected hitting time, (9), applies to all four cases. The

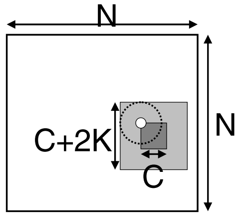


Fig. 4. Illustration of the expansion of the "footage" of community.

only difference is in the unit-time hitting probability, $P_{h,n}$ and $P_{h,c}$. The general expression for $P_{h,n}$ is

$$P_{h,n} = I(target\ in\ comm.\ in\ NMP) P_{move,l,n} 2K\overline{v_{l,n}}/C_n^2 + P_{move,r,n} 2K\overline{v_r}/N^2, \tag{11}$$

where $I(\cdot)$ is the indicator function. The expression for $P_{h,c}$ can be derived similarly.

B. Derivation of the Meeting Time

The derivation of the meeting time is similar to the hitting time detailed in the last section. In the following derivation, again we first arrive at the meeting probability between two nodes in a unit-time slice and the meeting probability of each time period. Then, in a very similar fashion to Theorem 4.4, we derive the meeting time by separating the sub-cases of meeting in the NMP or the CMP, and adding up the time components of whole periods and fraction of the last period in which the meeting event occurs.

The meeting time calculation heavily depends on the relative location of the communities of the involved nodes. Since nodes move within their corresponding communities more often than roaming out of the communities, it is obvious that two nodes with overlapping communities should meet each other much faster. We first derive the probability of nodes having overlapped communities, and then derive the meeting probabilities for both cases, overlapped or non-overlapped communities.

Lemma 4.5: For a specific type of time period, the communities of two nodes overlap with probability

$$P_x = \frac{(C+2K)^2}{N^2}. \tag{12}$$

Proof: As shown in Fig. 4, when a mobile node moves within its community, the area covered by the node (i.e., the area that could fall in the communication range of the node) actually extends out of the community by the transmission range of the node. The "footage" of the community is hence larger than C^2. We approximate this area by $(C + 2K)^2$, ignoring the small differences at the corners.

If the other node has its community chosen within this area, the meeting probability between the nodes would be much larger. Since each node selects its community at random within the simulation area, the probability that part of the footage of the community of node 1 is chosen as part of the community of node 2 is simply $\frac{(C+2K)^2}{N^2}$. ■

We now move on to derive the unit-time meeting probability. The meeting probability for nodes with overlapped communities and non-overlapped communities are derived separately.

Lemma 4.6: The unit-time meeting probability for nodes with non-overlapped communities is

$$\begin{aligned} P_{m,no\_ov} = & \frac{2K\overline{v_r} \times 2P_{move,r}(P_{pause,r} + P_{pause,l})}{N^2} \\ & + \frac{2K\overline{v_l} \times 2P_{move,l}P_{pause,r}}{C^2} \times \frac{C^2}{N^2} \\ & + \frac{2K\overline{v}\hat{v}((P_{move,r} + P_{move,l})^2 - P_{move,l}^2)}{N^2}, \end{aligned} \tag{13}$$

and the unit-time meeting probability for nodes with overlapped communities is

$$\begin{aligned} P_{m,ov} = & \frac{2K\overline{v_l}\hat{v}P_{move,l}^2}{C^2} + \frac{2K\overline{v_l} \times 2P_{move,l}P_{pause,l}}{C^2} \\ & + \frac{2K\overline{v_r} \times 2P_{move,r}(P_{pause,r} + P_{pause,l})}{N^2} \\ & + \frac{2K\overline{v_l} \times 2P_{move,l}P_{pause,r}}{C^2} \times \frac{C^2}{N^2} \\ & + \frac{2K\overline{v}\hat{v}((P_{move,r} + P_{move,l})^2 - P_{move,l}^2)}{N^2}. \end{aligned} \tag{14}$$

Proof: If the communities do not overlap, the nodes can only meet when at least one of them is out of the community (i.e., roaming). The first and the second terms in (13) correspond to the scenario when one node is moving and the other is not. In the first term, the moving node is in roaming state and the non-moving node can be in either local or roaming state. The moving node covers $2K\overline{v_r}$ new area each time unit. Since it performs a roaming movement, it meets with the other node with probability $2K\overline{v_r}/N^2$ as it does not have a priori knowledge about where the paused node is. In the second term the moving node performs a local movement and the paused node in roaming epoch happens to pause within the community of the moving node, which happens with probability C^2/N^2. Since the moving node moves locally, it meets with the other node with probability $2K\overline{v_l}/C^2$.

The third term corresponds to the scenario when both nodes are moving. We make use of the fact that when both nodes move according to the random direction model, the effective extra area covered can be captured by the multiplicative factor of relative speed, $\hat{v}$, which is 1.27 [16]. Note that the two nodes cannot meet if they both perform local movement, hence we have to multiply the meeting probability by the factor of $((P_{move,r} + P_{move,l})^2 - P_{move,l}^2)$ (i.e., at least one of them is moving in roaming epoch).

If the communities overlap, the nodes meet with higher probability when they both perform local movements. Here we make the simplifying assumption that the two communities are perfectly overlapped. As we will show in section IV-C, the theory is reasonably close to the simulation despite such simplification. The first two terms in (14) correspond to the scenario when both nodes are in local epochs. Under such scenario, the new area covered by a moving node contains the other node with probability $2K\overline{v_l}/C^2$. The first term captures the scenario when both nodes move locally and the second term captures the scenario when only one node moves, with similar reasonings as above. The remaining terms correspond to the scenario when at least one of the nodes is in roaming epoch. They are exactly the same as in the sub-cases with non-overlapped communities. ■

Following the unit-time meeting probability for the sub-cases of overlapped or non-overlapped communities in Lemma 4.6 and the probability for community overlap derived in Lemma 4.5, we have:

Corollary 4.7: The probability for at least one meeting event to occur during NMP and CMP, respectively, are

$$P_{M,n} = 1-P_{x,n}(1-P_{m,ov,n})^{T_n}-(1-P_{x,n})(1-P_{m,no\_ov,n})^{T_n}, \quad (15)$$

$$P_{M,c} = 1-P_{x,c}(1-P_{m,ov,c})^{T_c}-(1-P_{x,c})(1-P_{m,no\_ov,c})^{T_c}. \quad (16)$$

In an almost parallel fashion to Theorem 4.4, the expected meeting time can be calculated using the results in the Lemmas in this section.

Theorem 4.8: The expected meeting time is:

$$\begin{aligned}
MT &= Prob(meet\ in\ NMP)MT_{meet\ in\ NMP} \\
&\quad + Prob(meet\ in\ CMP)MT_{meet\ in\ CMP} \\
&= \frac{P_{M,n}}{Q}(T_n+T_c)(1/Q-1) \\
&\quad + \frac{P_{M,n}}{Q}\frac{1-(1+T_nP_{m,n})(1-P_{m,n})^{T_n}}{P_{m,n}(1-(1-P_{m,n})^{T_n})} \\
&\quad + \frac{(1-P_{M,n})P_{M,c}}{Q}(\frac{T_n+T_c}{Q}-T_c) \\
&\quad + \frac{(1-P_{M,n})P_{M,c}}{Q}\frac{1-(1+T_cP_{m,c})(1-P_{m,c})^{T_c}}{P_{m,c}(1-(1-P_{m,c})^{T_c})},
\end{aligned} \quad (17)$$

where $Q = P_{M,n} + P_{M,c} - P_{M,n}P_{M,c}$ is the probability for at least one meeting event to occur in one full cycle of NMP and CMP.

Proof: The proof is parallel to the proof of Theorem 4.4 and we omit it due to space constraints. ■

C. Validation of Theory with Simulations

In this section we compare the theoretical values of the hitting time and the meeting time to the corresponding simulation results, under various parameter settings. We summarize the parameters for tested models in Table II. Among them, model 1 is the case when the model behaves similar to the MIT WLAN trace (details in the next section). Model 2 is a scenario that communities have strong attraction. Model 3 is a scenario that communities are not very attractive and the nodes have equal probability to perform local and roaming movements in the NMP. Model 4 features large communities.

We perform simulations for the hitting and the meeting times for $50,000$ independent iterations and compare the average results with theoretical values derived from (5) and (17). Our discrete-time simulator is written in C, and nodes move as per descriptions in section III. To find out the hitting or the meeting time, we move the nodes in the simulator indefinitely until they hit the target or meet with each other, respectively. As shown in Fig. 5 for model 1, the simulation results and theoretical values are very close to each other. The relative error between theory and simulation results is between 4.62% to 10.38% for the hitting time and between 0.59% to 4.27% for the meeting time. Note that the simulator is not a repetition of theoretical calculations and hence the matching results validate each other. More details about the simulator and the codes can be found at [24].

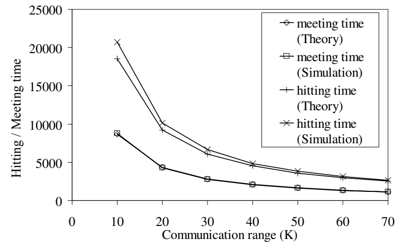
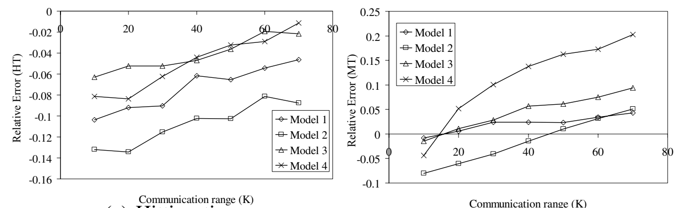


Fig. 5. Comparing the simulation results and theoretical values for the hitting and the meeting times (model 1).



(a) Hitting time. (b) Meeting time.

Fig. 6. Relative error between theory and simulation values.

We show the relative errors between the theoretical values and the simulation results for other models in Fig. 6. The relative error is calculated as $Error = (Theory - Simulation)/Simulation$. Hence a positive error indicates the theoretical value is larger than the simulation result, while a negative error indicates the converse. From Fig. 6 we see that for all four models, the relative errors are within acceptable range. The absolute values for the error are within 15% for the hitting time and within 20% for the meeting time. For 80% of the tested parameter sets, the error is below 10%. These results display the accuracy of our theory under a wide parameter settings, especially when the communities are small compared to the simulation area.

V. Matching Real and Synthetic Traces

We have shown in the previous section that the hitting time and the meeting time can be derived for the proposed time-variant community mobility model. In this section we in turn show that in addition to theoretical tractability, the model is also generic enough to be fine-tuned and display matching mobility characteristics with the WLAN traces, in terms of both *skewed location visiting preferences* and *periodical re-appearance at the same location*.

To tally *location preferences* and *re-appearance probability* for the continuous time-variant community mobility model, we first create a synthetic WLAN trace from the mobility model.

TABLE II

PARAMETERS FOR THE SCENARIOS IN THE SIMULATION

Model name	Description	N	C_n	C_c	v_{max}, v_{min}	$T_{max,n}$	$T_{max,c}$	$\overline{L_r}$	$\overline{L_l}$	$p_{l,n}$	$p_{r,n}$	$p_{l,c}$	$p_{r,c}$	T_n	T_c
Model 1	Match with the MIT trace	1000	100	100	$15, 5$	100	50	520	80	0.5	0.2	0.8	0.2	5760	2880
Model 2	Highly attractive communities	1000	200	50	$15, 5$	100	200	520	52	0.6	0.3	0.8	0.1	3000	2000
Model 3	Not attractive communities	1000	100	100	$15, 5$	50	200	800	80	0.5	0.5	0.6	0.3	2000	1000
Model 4	Large-size communities	1000	200	250	$15, 5$	50	100	800	200	0.7	0.3	0.8	0.1	2000	1000

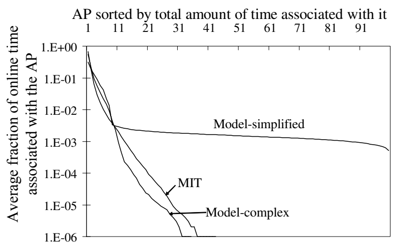


Fig. 7. Matching *location visiting preferences* of the synthetic traces to the WLAN trace.

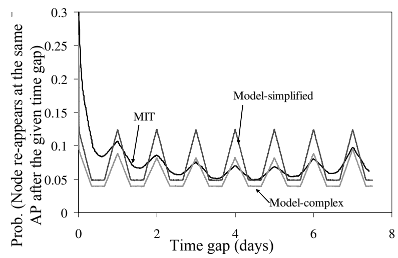


Fig. 8. Matching *periodical re-appearance* of the synthetic traces to the WLAN trace.

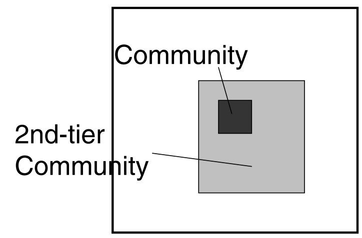


Fig. 9. Illustration of multi-tier community (3-tiers).

We do so by introducing the notion of *virtual* access points in the simulation area and defining when the mobile nodes are associated with these *virtual* APs. Note that the behavior of mobile nodes do not change from the description in section III by the introduction of these *virtual* APs. They are introduced solely for the purpose of tallying mobility characteristics. We divide the whole simulation area into 100 regular, equal-sized grids. Each grid is covered by one *virtual* AP. Since devices are usually turned off when users start moving them in the real WLANs, we make a similar assumption that the mobile nodes are considered associated with the *virtual* AP in the current grid they reside in only when they are not moving. We then tally the mobility characteristics from the synthetic trace by the same way as we treat real WLAN traces.

We use the MIT WLAN trace (presented in [9]) as an example to show that our model, with parameters carefully adjusted, could generate a synthetic trace that mimics the behavior of real WLAN traces. We also achieved good matching with the USC[22] or the Dartmouth[11] traces, but do not show it here due to space constraints. In Fig. 7, we show the *skewed location visiting preferences* property. In both the MIT trace and the synthetic trace from our simple model (the curve labeled as *model-simplified*, using parameters of Model 1 in Table II), the nodes stay within their favorite locations (i.e., the communities) for high probability. However, for the WLAN trace the curve shows a fast decaying tail, while for our model the tail of the curve levels off. This is because the node roams within the whole simulation area and pauses at any location with equal probability in roaming epochs. We will show a method to extend our simple model to resolve this issue below.

In Fig. 8, we show the *periodical re-appearance* property. In the MIT trace, the nodes re-appear at the same AP with higher probability if the time gap between the considered time instants is an integer multiple of days. As shown in the figure (*model-simplified*), we can set the time periods, such that the CMP captures the working hours (i.e., nodes go to office with very high probability) and the NMP models off-work hours. The synthetic trace mimics the peaks of the re-appearance probability. However, since we have only two types of time periods in the model, the peaks repeat itself with the same value, while in the WLAN trace we also observe *weekly* periodicity: the probability for nodes to re-appear at the same AP is higher if the time gap is seven days.

Our simplified model, as described in section III, makes the theoretical derivation of the hitting time and the meeting time cleaner to handle. But, on the other hand, it may be insufficient to fully capture complex user behavior, as shown by comparing curves of *model-simplified* and *MIT* in Fig. 7 and 8. To resolve the mismatch, we introduce more fine-grained model both in space and time. In the space domain, we introduce multi-tier communities, as illustrated in Fig. 9 with a three-tier example. The node visits each tier of the community with decreasing probability. This extension is suggested by the intuition that we move locally close to homes and offices more than go far from these places. This way, we allow a smooth decaying tail rather than the horizontal tail in the simple model. In the time domain, we add more time periods with different mobility parameters. These additional time periods can be used to capture either more detailed mobility behavior for different time periods within a day or differences of mobility in weekdays and weekends. Finally, we use a more complex model with six-tier communities and three distinct time periods to model weekday work hour, weekend day time, and night time separately. The synthetic trace derived from this model matches well with the MIT trace in terms of both *location visiting preferences* and *periodical re-appearance*. The curves labeled as *Model-complex* in Fig. 7 and Fig. 8 show the mobility characteristics for this complex model.

Fine-tuning our model to match with the mobile node behavior in the WLAN trace is more of a trial and error process. We first adjust the attraction from each tier of the community (i.e., the likelihood of having an epoch in the tier) and the pause times of the node in each tier to shape a decaying tail of *location visiting preference*. This is typically achieved by a power-law decaying attraction and pause time at each community tier. The time ratio nodes spend on moving and pausing is then adjusted to change the peaks in the curve of *periodical re-appearance*. The key point is, once these properties are collected from the targeted environment, our model provides a flexible platform which can be tuned to match with the desired scenario.

Note that the theory derivations of the hitting time and the meeting time with these extensions are technically feasible by simply adding more sub-cases. For the hitting time, we split the sub-cases that the target falls in each tier of the community in (5). For the meeting time, we need to add more sub-cases that the communities of two nodes overlap with various degrees (i.e., up to a certain tier), in addition to (13) and (14). Due to lack of space, we do not present this analysis in detail here.

Currently we assume that each node randomly picks its own community from the simulation area. However, in realistic settings there should be points of common interest in the environment, at least during certain time periods (e.g., restaurants on campus during lunch time, hotels in an amusement park during nights). To capture such common attraction points, we could use an alternative way to set up communities: We can fix the locations of multiple communities serving as known point-of-interest in the targeted scenario, and let the nodes choose one of them at random with given probabilities during concentration periods. This set up captures that nodes will meet at these attraction points with higher probability during CMP. The only change needed in the theory is the probability of community overlap in the derivation of the meeting time. Instead of using (12), the probability of community overlap, P_x, is simply the probability of choosing the same attraction point to visit during CMP, which can be easily derived.

VI. Conclusion and Future Work

We have proposed a *time-variant community mobility model* for wireless mobile networks. Our model preserves two characteristics of mobility observed in current WLANs, namely *skewed location visiting preferences* and *periodical re-appearance at the same location*. In addition to providing realistic mobility patterns, the proposed model can be mathematically analyzed to derive two quantities of interest, the *hitting time* and the *meeting time*. These two quantities are closely related to the routing performance in delay-tolerant networks or mobility-assisted packet forwarding. Through extensive simulation studies, we have verified the accuracy of our theory - the error is always below 20% and under 10% for most cases.

In the future we would like to extend our theory to the scenario in which nodes display more heterogeneous behavior. Nodes may have different set of parameters and the time period structure can be different for the nodes. With these extensions, we will have a mobility model to describe an environment including users with diverse mobility characteristics. We believe such a model is a very important building block for understanding protocol performances in real-life settings.

Acknowledgement

The authors would like to thank the anonymous reviewers for providing valuable comments. This work is supported by NSF awards CNS-0520017 and career award 0134650.

References

[1] Delay tolerant networking research group. http://www.dtnrg.org.
[2] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," In Proceedings of ACM SIGCOMM, Aug. 2004.
[3] 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.
[4] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Single-copy routing in intermittently connected mobile networks," In Proceedings of IEEE SECON, Oct. 2004.
[5] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: Efficient routing in intermittently connected mobile networks," In Proceedings of ACM SIGCOMM workshop on Delay Tolerant Networking (WDTN), Aug. 2005.
[6] P. Gupta and P. Kumar, "Capacity of wireless networks," Transactions on Information Theory, 46(2), 2000.
[7] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," IEEE/ACM Transactions on Networking, 10(4), 2002.
[8] S. Ross, "Introduction to Probability Models," 8th edition, Academic Press, published Dec. 2002.
[9] M. Balazinska and P. Castro, "Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network," In Proceedings of MOBISYS 2003, May 2003.
[10] M. McNett and G. Voelker, "Access and mobility of wireless PDA users," ACM SIGMOBILE Mobile Computing and Communications Review, v.7 n.4, Oct. 2003.
[11] T. Henderson, D. Kotz and I. Abyzov, "The Changing Usage of a Mature Campus-wide Wireless Network," In Proceedings of ACM MOBICOM, Sep. 2004.
[12] W. Hsu and A. Helmy, "On Important Aspects of Modeling User Associations in Wireless LAN Traces," In Proceedings of the Second International Workshop On Wireless Network Measurement, Apr. 2006.
[13] C. Tuduce and T. Gross, "A Mobility Model Based on WLAN Traces and its Validation," In Proceedings of IEEE INFOCOM, Mar. 2005.
[14] R. Jain, D. Lelescu, M. Balakrishnan, "Model T: An Empirical Model for User Registration Patterns in a Campus Wireless LAN," In Proceedings of ACM MOBICOM, Aug. 2005.
[15] D. Lelescu, U. C. Kozat, R. Jain, and M. Balakrishnan, "Model T++: An Empirical Joint Space-Time Registration Model," In Proceedings of ACM MOBIHOC, May 2006.
[16] T. Spyropoulos and K. Psounis, "Performance Analysis of Mobility-assisted Routing," In Proceedings of ACM MOBIHOC, May 2006.
[17] R. Groenevelt, P. Nain, and G. Koole, "The Message Delay in Mobile Ad Hoc Networks," In Proceedings of PERFORMANCE, Oct. 2005.
[18] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," In Proceedings of IEEE INFOCOM, Apr. 2006.
[19] M. Musolesi and C. Mascolo, "A Community Based Mobility Model for Ad Hoc Network Research," In Proceedings of the Second International Workshop on Multi-hop Ad Hoc Networks (REALMAN), May 2006.
[20] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms," In Proceedings of IEEE INFOCOM, Apr. 2006.
[21] J.Su, A. Chin, A. Popivanova, A. Goel, and E. de Lara, "User Mobility for Opportunistic Ad-Hoc Networking," In Proceedings of the 6th IEEE Workshop on Mobile Computing Systems and Applications, Dec. 2004.
[22] MobiLib: Community-wide Library of Mobility and Wireless Networks Measurements. http://nile.usc.edu/MobiLib.
[23] CRAWDAD: A Community Resource for Archiving Wireless Data At Dartmouth. http://crawdad.cs.dartmouth.edu/index.php.
[24] Simulation codes used in this work and its detailed description are available at http://nile.cise.ufl.edu/~weijenhs/TVC_model