

Users Mobility Models for Opportunistic Networks: the Role of Physical Locations

Chiara Boldrini, Marco Conti, and Andrea Passarella
CNR-IIT, Via G. Moruzzi, 1 - 56124 Pisa, Italy
{c.boldrini,m.conti,a.passarella}@iit.cnr.it

Abstract—Opportunistic networks are pervasive networks built exclusively by users’ devices. With respect to legacy multi-hop ad hoc networks (MANETs), opportunistic networks are much more resilient to disconnections, long network partitions, and dynamic reconfigurations. Therefore, they are suitable for challenged scenarios, such as disaster areas, or rural environments. While in MANET nodes’ mobility is an issue, in opportunistic networks mobility is an opportunity for communication, therefore mobility models are a key component in opportunistic networks’ research. Improved (with respect to MANETs) users mobility models are rapidly becoming reference tools for studying opportunistic networks, thanks to their accuracy in matching statistical features observed in real traces. A very interesting approach is defining mobility patterns based on *social relationships* between users. In this paper we argue that, while being a fundamental building block, social relationships alone cannot be the only driver for the mobility process. Users movements are also driven by the attraction of *physical locations* (e.g., the house, the working place) on users. We analyse from this standpoint the Community-based Mobility Model (CMM), which is the reference point for social-aware community models. We show in detail that CMM is not able to capture locations’ attractions on users. Therefore, we propose the Home-cell Community-based Mobility Model (HCMM), which retains the social-aware features of CMM and, in addition, permits to easily control the relationships between users and their preferred physical locations.

I. INTRODUCTION

Opportunistic networks are wireless mobile networks built entirely on users’ devices [1]. With respect to legacy MANETs, no end-to-end path is assumed between the sender and the receiver at the time of data generation. Paths are built dynamically on-the-fly while messages progress towards the destination, opportunistically exploiting contacts between nodes. Therefore, they are much more tolerant than MANETs to network partitions (also lasting for long times), and reconfigurations, and are thus particularly suitable for challenged scenarios such as disaster, emergency and rural networks.

Mobility modelling for opportunistic networks is a hot topic in the research community. Opportunistic networks actually *exploit* users’ mobility to bridge disconnections and partitions [2]. Therefore, it is of paramount importance to identify realistic mobility models, both to drive the protocols’ design, and to provide sensible performance results. In the last few years, there has been an increasing effort aimed at reconsidering the MANET mobility models [3] for opportunistic networking scenarios. There is general agreement

on the fact that popular models used in MANET research (e.g., the random waypoint model) generate quite unrealistic users’ behaviour (e.g., [4], [5]). To address this issue, mobility models are reconsidered or re-designed based on real users’ mobility traces available to the community (e.g., through CRAWDAD).

Several proposal ([5], [6], [7]) exploit WLAN association traces to derive users’ association profiles and, based on these, mobility models. The resulting models are very good in capturing the fact that physical locations (WLAN hotspots in this case) exert attraction on users. The work in [8] takes this idea one step further, and provides mobility models in which general physical locations (not necessarily WLAN access points) exert attractions on users. Finally, authors of [9] explain WLAN association traces with sociological-inspired concepts, noticing that periodic association patterns follow sociological orbits, defined by the users’ social behaviour. Exploiting this remark, they provide a user-centric model (rather than an “AP”-centric one as in the previous works). This body of work is based on the fundamental observation that users are attracted by particular physical locations, in which they tend to preferentially spend their time. The limit we see in this approach is the fact that it does not explain the mechanisms resulting in the modelled mobility patterns. Therefore, it is not clear if the resulting models are applicable to networking scenarios other than the ones used for the initial observations (most notably, if they are applicable to opportunistic networks too).

Exploiting the social behaviour of users to define the basic mechanisms of users’ movements is a very interesting direction. To the best of our knowledge, the most advanced proposal of this class is the work in [4], where authors define the Community-based Mobility Model (described in Section II-A). The most interesting feature of CMM is the leveraging of social network theories and models [10] to define users’ movements. Besides matching well real users’ mobility traces [4], this approach sheds light on the features of users’ social behaviour that result in the mobility features observed in real traces.

Despite these nice properties, in this work we show that the original CMM proposal is not able to capture the attraction exerted on users by physical locations. Specifically, we show in Sections II-B and II-C that CMM shows a *gregarious* behaviour, such that all users in a community tend to follow the first user that moves outside the physical location where

the community is located. The gregarious behaviour does not represent significant scenarios (e.g., working places), where users roam around preferred physical places, besides being influenced by social relationships between each other. To address this issue, we propose the *Home-cell Community-based Mobility Model (HCMM)*, which joins the concepts of CMM (for modelling social relationships between users) with the concept of defining preferential locations in which users tend to spend most of their time. Therefore, HCMM is a first step towards joining together the two promising mobility modelling approaches discussed above. After describing HCMM (in Section III) we show that it still matches characteristic features of real traces. Furthermore, we highlight that, unlike CMM, it provides very simple knobs to control the time spent by users in their preferred physical locations (Section III-A).

II. MODELLING BASED ONLY ON SOCIAL ATTRACTION

A. Community-based Mobility Models

In CMM [4] each node belongs to a social community (group). Nodes that are in the same social community are called *friends*, while nodes in different communities are called *non-friends*. Relationships between nodes are modelled through social links (each link has an associated weight). At the system start-up all friends have a link to each other. Links between non-friends are set up according to the *rewiring probability* (p_r) parameter. For each node, each link (towards a friend node) is rewired (i.e., it is moved to connect the node and a non-friend node) with probability p_r . If the link has to be rewired, the target non-friend node is randomly selected according to a uniform distribution.

Social links are used in CMM to drive node movements. Nodes move in a grid, and each community is initially randomly placed in a cell of the grid. Each node movements can be seen as a sequence of steps. At the beginning of each step, the node selects the two components that define the next movement: the cell towards which to move, and the “goal” within the target cell. The target cell is selected according to the social attraction exerted by each cell on the node. Attraction is measured as the sum of the links’ weights between the node and the nodes currently moving in or towards the cell. The target cell is chosen based on the probabilities defined by cells’ attraction (i.e., if a_j is the attraction of cell j , then the probability of selecting that cell is $a_j / \sum_j a_j$). Finally, the “goal” within the target cell is selected according to a uniform distribution. Speed is also selected accordingly to a uniform distribution within a user-specified range¹.

In the following we will study how a tagged node k that goes out of its starting cell influences the movements of the other nodes of its community. Without loss of generality, we consider a simple scenario with two communities only, placed in two distinct cells. At the system startup, n nodes are placed in node k ’s cell, and $f \cdot n$ nodes in the other cell (hereafter, node k ’s cell is denoted as starting cell, and the

other cell as destination cell), f being a parameter greater than 0. All nodes but node k have only social links within their community, while node k has also links with nodes in the other community. Due to these links, node k , at some point in time, moves to the destination cell. To investigate the gregarious behaviour, we compute the *remaining probability* (P_{rem}), defined as the probability of *no other* member of node k ’s community to move towards the destination cell. When P_{rem} approaches 0, at least one node in the starting cell follows node k . As will be clear from the following analysis, this may generate an avalanche effect such that all nodes in node k ’s community follow node k in the destination cell, thus revealing the gregarious behaviour.

We choose to study the case of a single node (k) having links outside its community, because, as we will show analytically, it represents the weaker condition for the gregarious behaviour to take place. Therefore, the P_{rem} formula computed in the following section is actually an upper bound of the remaining probability achieved in the general case.

B. Gregarious Behaviour in CMM: Analytical Model

In CMM each cell $S_{p,q}$ in the grid (i.e., with generic coordinates p, q) exerts on a node i an attraction equal to

$$SA_{p,q}^{(i)} = \frac{\sum_{j \in C_{S_{p,q}}} w_{i,j}}{c} \quad (1)$$

with $C_{S_{p,q}}$ being the set of nodes associated to cell $S_{p,q}$ (i.e., all nodes whose *current goal* is within that cell²), c the number of nodes associated to $S_{p,q}$ and $w_{i,j}$ the weight of the link (relationship) between node i and node j . To select the destination cell, each node i computes its attractions towards each cell (according to Equation 1), and then selects the cell $S_{p,q}$ with probability $P(s = S_{p,q}) = SA_{p,q}^{(i)} / \sum_{j=1}^{p \times q} SA_j^{(i)}$.

After node k starts moving outside the starting cell, each node i of its community (i.e $C_k = C_i$) is attracted by the destination cell with a social attractions equal to

$$SA_{(dest)}^{(i)} = w_{i,k} / (fn + 1) \quad (2)$$

and by the starting cell with a social attraction equal to

$$SA_{(start)}^{(i)} = \sum_{j \in C_{start}}^{n-1} w_{i,j} / (n - 1). \quad (3)$$

Equations 2 and 3 already show three interesting features. Firstly, the attraction of node k on other nodes in the same community depends not only on the strength of the social relationships within the community ($w_{i,k}$), but also on the number of nodes associated to the destination cell ($fn + 1$). The more the nodes associated to the destination cell, the less the attraction that k exerts on the other nodes. Secondly, it can already be shown that we are considering a worst-case scenario for the gregarious behaviour to occur. If several nodes had social relationships outside their community, there would

¹CMM also includes further mechanisms, such as *reconfigurations*, whose effect is however orthogonal to what we focus on in this paper. The interested reader is referred to [4].

²Note that, based on this definition, a node is associated to the destination cell as soon as it *selects* that cell as its next goal.

be more nodes going outside the starting cell. The attraction exerted by the destination cell (or the joint attraction exerted by destination cells, more in general) on each remaining node will be greater than $SA_{(dest)}^{(i)}$, thus increasing the probability of at least one such node to go out the starting cell. Finally, it is also easy to understand that considering the remaining probability allows us to highlight a possible avalanche effect. When the remaining probability is close to 0, at least one node follows node k outside the starting cell. This increases the attraction exerted by node k 's destination cell on nodes in the starting cell, thus further increasing the probability of other members of node k 's community to move towards the destination cell.

Going back to our reference scenario, the probability that a generic node i in node k 's community goes out of the starting cell is equal to

$$P_{out_i} = \frac{w_{i,k}/(fn+1)}{w_{i,k}/(fn+1) + \sum_{j=1}^{n-1} w_{i,j}/(n-1)} \quad (4)$$

and thus $1 - P_{out_i}$ represents the probability that node i remains in the starting cell at the next step. For the sake of simplicity, we assume $w_{i,k} = w_{j,k} = w_k, \forall i, j \mid C_i = C_j$. In addition, we approximate $\bar{w} \simeq \sum_j w_{i,j}/(n-1)$, where \bar{w} denotes the average of the distribution of weights between nodes of the same community. Equation 4 thus becomes as follows:

$$P_{out} = \frac{w_k/(fn+1)}{w_k/(fn+1) + \bar{w}}. \quad (5)$$

Note that a special case is represented by $n = 2$. In this case, $SA_{(start)}^{(i)}$ is clearly 0 (just one node is in the starting cell after node k leaves), and therefore P_{out} is equal to 1.

The longer node k is associated to the destination cell, the longer other nodes have opportunities of following it. In the worst case (from the gregarious behaviour standpoint) node k comes back to the starting cell right after reaching the goal in the destination cell (this corresponds to the minimum possible time during which node k is associated to the destination cell). The number of times a node i (in the same community of node k) runs the next-goal selection algorithm while node k is associated to the destination cell is thus equal to the ratio between the average time it takes for node k to reach the goal in the destination cell ($\bar{T}^{(out)}$) and the average time it takes for node i to complete a random movement inside the starting cell ($\bar{T}^{(in)}$). Under the assumption of a $n \times n$ grid and square cells, the average length of a random movement \bar{D} inside the grid can be assumed to be proportional to the average length of a random movement \bar{d} inside a single cell, with proportionality constant l (equal to the ratio between the edge of the grid and the edge of the cell). If we define \bar{V} as the average speed at which nodes move, then $\frac{\bar{T}^{(out)}}{\bar{T}^{(in)}} = \frac{\bar{D}}{\bar{V}} \cdot \frac{\bar{V}}{\bar{d}} = \frac{\bar{D}}{\bar{d}} = l$ holds true.

As there are $n - 1$ nodes in the starting cell (after node k 's departure), the remaining probability can be computed as:

$$P_{rem} = [(1 - P_{out})^l]^{n-1} = \left[\left(1 - \frac{w_k/(fn+1)}{w_k/(fn+1) + \bar{w}} \right)^l \right]^{n-1}, \quad (6)$$

Parameter	Value
n	10
l	5
\bar{w}	0.5
w_k	0.5
f	1

Fig. 1. Parameters' default values

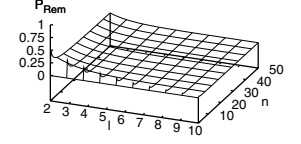


Fig. 2. P_{rem} as a function of n and l .

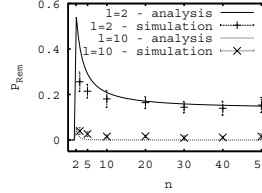


Fig. 3. Validation of the analytical model (P_{rem} as a function of n).

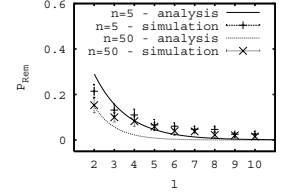


Fig. 4. Validation of the analytical model (P_{rem} as a function of l).

where l is the average number of times each node in the starting cell selects a new destination while node k is associated with the destination cell, $1 - P_{out}$ is the probability of each node to select the starting cell for the next step, and $n - 1$ is the number of nodes in the starting cell after node k departure.

C. Gregarious Behaviour in CMM: Quantitative Analysis

In this section we analyse the impact of the mobility model parameters on the remaining probability derived in Equation 6. Unless otherwise stated, parameters are set according to Figure 1. Note that in the following we consider two alternatives for assigning relationships between nodes (weights) before rewiring. In the first one, weights are uniformly distributed between 0 (not included) and 1. In the second one, weights between friends are uniformly distributed between $1 - threshold$ and 1, with $threshold$ being a configurable parameter ranging from 0 to 1. The second alternative is the one actually implemented in CMM because it allows for controlling the average weight within a community. Without the threshold, the average weight between friend nodes (\bar{w}) is stuck at 0.5, while, with the threshold-based approach, \bar{w} is equal to $1 - threshold/2$, and it thus ranges between 0.5 and 1.

Figure 2 illustrates the P_{rem} dependence on n (the number of the nodes of k 's community), and l (the ratio between the movements duration outside and within a community). Figures 3 and 4 compares selected curves from the plot in Figure 2 (thus derived through the model) with simulation results (confidence intervals have 90% confidence level)³. This comparison shows that the analytical model is accurate, particularly starting from medium values of n and l .

For small values of l , the grid has few cells and the duration of k 's movement outside the starting cell is not so different from the duration of nodes' random movement inside a cell. Thus, a generic node i has not many opportunities of going outside the starting cell, because node k is associated with the destination cell only for a relatively small amount of time. The trend highlighted in Figure 2 generally holds true

³Similar curves have been derived also with respect to other parameters analysed in the following (not reported here due to space reasons).

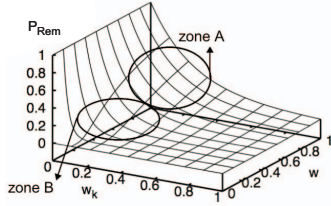


Fig. 5. P_{rem} as a function of w_k and \bar{w} .

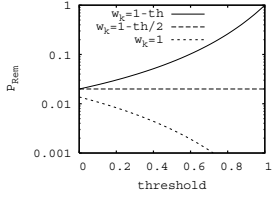


Fig. 6. P_{rem} as a function of the *threshold* (uniform distribution of the weights).

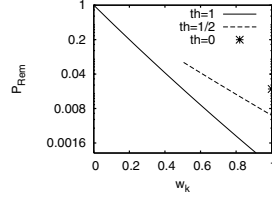


Fig. 7. P_{rem} as a function of w_k (uniform distribution of the weights).

when considering the impact of l , irrespectively of the other parameters' configurations. Therefore, we will not analyse the impact of l further on.

To better understand the behaviour with respect to n , let us rewrite Equation 6, by recalling that $\bar{w} = w_k$. It is easy to show that Equation 6 becomes $P_{rem} = \left[1 - (1/n + 2)\right]^{n-1}$. The remaining probability of a *single* node $(1 - (1/n + 2))$ increases with n , because a large n corresponds to a ‘‘heavy’’ community, that exerts a strong attraction on its members. However, as the number of nodes increases, it is more and more difficult that *all* nodes remain in starting cell. The joint effect (shown in Figure 2) is that P_{rem} is significantly greater than 0 only for small values of n .

In Figures 5, 6, and 7 we analyse how the strengths of social relationships (modelled through the parameters w_k and \bar{w}) impact on the remaining probability. Specifically, to have a complete view across the whole parameters' space, in Figures 5 we explore the whole range of possible values for the parameters, irrespectively of the particular distribution used to actually assign weights. Instead, Figures 6 and 7 show remaining probability curves obtained when the weights are assigned according to the uniform distribution considered in the original CMM model, by also exploiting the *threshold* parameter. Specifically, Figure 6 shows three curves, corresponding to the lowest, the average and the highest value of w_k . In Figure 7 curves start at the minimum w_k value, i.e., $w_k = 1 - \text{threshold}$.

In Figure 5 note that, when w_k is small, i.e. when node k exerts a weak attraction on the nodes of the starting cell, the main contribution to P_{rem} is due to \bar{w} (i.e., the attraction between nodes in the starting cell): if \bar{w} is high, nodes are more likely to remain in their current cell. However, as soon as w_k increases beyond 0.5, P_{rem} drops approximately to 0, regardless of the weights between nodes of the same cell. This is an important results, because it highlights that the gregarious behavior is a constant in CMM model for a large range of (w_k, \bar{w}) values. Remaining probabilities significantly greater than 0 are achieved only in the region (zone A in the Figure) with high attraction in the starting cell (high \bar{w}), and low at-

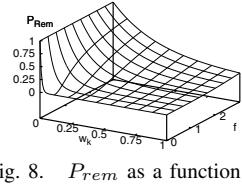


Fig. 8. P_{rem} as a function of w_k and f .

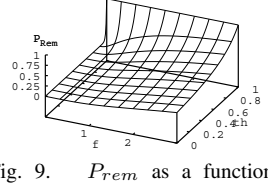


Fig. 9. P_{rem} as a function of f and *threshold*, in the worst-case scenario ($w_k = 1 - \text{threshold}$).

traction of the tagged node (low w_k). Figures 6 and 7 show that this combination cannot be achieved by a uniform distribution of the weights (as in the CMM model), even by controlling the average value through the *threshold* parameters. Non-negligible remaining probabilities are exclusively achieved for low w_k values (e.g., curve $w_k = 1 - th$ in Figure 6), when the *threshold* is high. As the average weight (\bar{w}) is equal to $1 - \text{threshold}/2$, in this configuration \bar{w} turns out to be low. The region with maximum remaining probability in the (w_k, \bar{w}) space where CMM can actually operate is thus zone B in Figure 5, while it can never operate in zone A.

In this first set of plots we considered f equal to 1 (i.e., the same number of nodes in the starting and destination cells). This choice is coherent with the original implementation of the CMM model, in which the number of nodes in each cell at the system start up is almost the same. In real social networks, however, non homogeneous communities are quite common. In the following we thus investigate the gregarious behavior with non homogeneously populated communities. Equation 2 highlights how f impacts on the social attraction of the destination cell. When the destination cell is more populated than the starting cell ($f > 1$), node k 's attraction is smoothed out, while, when it is less populated ($f < 1$), the attraction is amplified.

Figures 8 and 9 basically confirm the behaviour highlighted before. Even when the destination cell is more populated ($f > 1$), non-negligible remaining probabilities are achieved when the attraction of the tagged node is rather weak (low w_k values). Clearly, the greater the destination cell's population, the lower the attraction it exerts on members of node k 's community, the greater the remaining probability. Specifically, we have found that the remaining probability is always greater than 0.5 as f increases beyond 10. However, note that such high values for f are not that sensible, as $f = 10$ means that the destination community has 10 times more members than the starting community. Finally, Figure 9 shows that, for more sensible ranges of f , the only chance of having high remaining probability is a very high *threshold*, joint with the minimum attraction of the tagged node ($w_k = 1 - th$), and a densely populated destination cell ($f > 1$). Note that when the *threshold* approaches 1, the attraction of the tagged node becomes negligible.

The results shown in this section clearly highlight that the gregarious behaviour is a characteristic feature of the original Community-based Mobility Model. As discussed in Section I, this behaviour does not represent a class of typical scenarios, in which users move based on social relationships, but also on attractions exerted by the *physical places* in which these

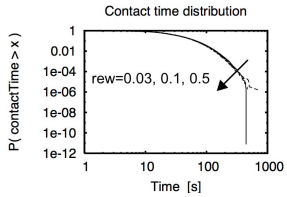


Fig. 10. Contact time CCDF

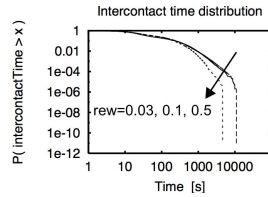


Fig. 11. Intercontact time CCDF

social relationships usually take place (e.g., working place). As this class of scenarios is quite large, the following section proposes a simple modification to the original CMM to address this point.

III. HOME-CELL COMMUNITY-BASED MOBILITY MODEL

In the Home-cell Community-based Mobility Model (HCMM), each node is attracted by its *home cell* (i.e. the cell to which its community is assigned after a reconfiguration), based on the social attraction exerted on that node by all other nodes that are part of its group, *irrespective* of the current physical location of those nodes. In a sense, the social links between nodes of the same group are translated in a global attraction exerted by the group home cell. Similarly, the social attraction towards an external cell is evaluated based on the social relationships with nodes having their home in that cell. When a node is in its home cell, the cell for the next movement is selected as in CMM. However, after a node reaches a cell which is not its home, it stays in the external cell with a given probability (p_e) for the next movement, and goes back home with probability $1 - p_e$. Therefore, it roams in the foreign cell for an average number of steps equal to $p_e / (1 - p_e)$. HCMM allows us to model a kind of scenario in which nodes are attracted towards a place (e.g., their office building) in which usually people of their group roam. Nodes are also attracted outside that place because of social relationships between groups, and spend some time in the foreign groups before heading back home.

In order to prove that HCMM generates realistic movement patterns, we analysed two properties: the *contact duration*, which is the time interval in which *any* two devices are in radio range, and *intercontact time*, defined as the time interval between two consecutive contacts between *any* two nodes. In real traces, the probability distribution of these figures shows a heavy tail [11]. For validating HCMM we ran a simulation with 30 nodes divided into three groups varying the rewiring parameter. Figure 10 and 11 show that HCMM still generates heavy-tail distributions for contact and inter-contact times, which is required to model realistic humans movements (the dependence of these figures on the rewiring probability is discussed in [12]). In Figure 10 all curves overlap, i.e. the contact duration doesn't depend on the value of the rewiring parameter. This is because we consider contacts between *any* two nodes, that are independent from the time a node remains in its group. In contrast, Figure 11 highlights the dependency of inter-contact times from the time a node spends in the same group. In fact, with increasing rewiring, nodes gets more and more mixed together and the time between two consecutive contacts tend to be smaller.

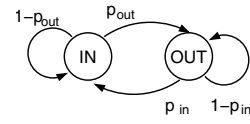


Fig. 12. Node's status in HCMM and CMM.

A. HCMM vs. CMM: Controlling Node Positions

In this section we compare HCMM and CMM. Specifically, we highlight the fact that HCMM allows for a fine control of the physical locations around which users' roam, while CMM does not provide any simple control parameter on this. To this end, we generalise the analytical model presented in Section II-B. The goal of the model we present hereafter is to provide closed formulas for the average time spent by any node inside and outside the starting cell (home cell in HCMM). For ease of presentation, we still assume to have just two cells, even though the destination cell can jointly represent all cells other than the starting cell. We assume that *all* links can be rewired at the system startup (with probability p_r). Therefore, we don't assume any difference between a tagged node (node k) and the other nodes anymore. We also don't focus anymore on the event of a particular node exiting the starting cell.

In HCMM and in CMM the status of each node can be represented with a 2-state discrete Markov chain as in Figure 12, where "IN" means the node is in the starting cell, and "OUT" means it is outside the starting cell. The difference between HCMM and CMM lies in the expressions of p_{in} and p_{out} , that we will derive at the end of this section. Otherwise, the analysis of the average time spent in the IN and OUT states is common to CMM and HCMM.

First of all, it is straightforward deriving the stationary distributions, $\pi_{in} = p_{in} / (p_{in} + p_{out})$, and $\pi_{out} = p_{out} / (p_{in} + p_{out})$. The average time spent in the IN and OUT states can be computed via the conditioned probabilities, as follows:

$$\begin{cases} E[T_{in}] &= \pi_{out} p_{in} \cdot E[T_{in} | E_{IN}] \\ E[T_{out}] &= \pi_{in} p_{out} \cdot E[T_{out} | E_{OUT}] \end{cases} \quad (7)$$

where E_{IN} and E_{OUT} denote the events "the node enters the IN state" and "the node enters the OUT state", respectively, while $\pi_{out} p_{in}$ and $\pi_{in} p_{out}$ are the probabilities of these events. By recalling that i) the number of steps spent in each state is distributed according to a geometric law, ii) the duration of each step both in the IN and OUT state can be approximated with $\bar{T}^{(in)}$, and iii) the duration of the transitions between the states can be approximated with $\bar{T}^{(out)}$ (see Section II-B), we can compute closed form expressions for $E[T_{in}]$ and $E[T_{out}]$ as follows:

$$\begin{cases} E[T_{in}] &= \frac{p_{in}(1-p_{out})}{p_{in}+p_{out}} \cdot \bar{T}^{(in)} \\ E[T_{out}] &= \frac{p_{out}(1-p_{in})}{p_{in}+p_{out}} \cdot \bar{T}^{(in)} + \frac{p_{in}p_{out}}{p_{in}+p_{out}} \cdot 2\bar{T}^{(out)} \end{cases} \quad (8)$$

To specialise Equation 8 to HCMM and CMM we have to compute the transition probabilities of the corresponding Markov chains, hereafter referred to as $p_{out}^{(H)}$ and $p_{in}^{(H)}$, and $p_{out}^{(C)}$ and $p_{in}^{(C)}$ respectively. By definition, $p_{in}^{(H)}$ is equal to $1 - p_e$. For the other parameters, we can use the following line of reasoning, common to HCMM and CMM. To compute

p_{out} , we should focus on a node *inside* the starting cell, and compute the attractions of the starting and destination cells. To compute p_{in} , we should compute the attractions of the starting and destination cells on a node *outside* the starting cell. Then, p_{in} and p_{out} can be computed as follows:

$$\begin{cases} p_{in} &= \frac{SA_{start}^{(out)}}{SA_{dest}^{(out)} + SA_{start}^{(out)}} \\ p_{out} &= \frac{SA_{dest}^{(in)}}{SA_{dest}^{(in)} + SA_{start}^{(in)}} \end{cases} \quad (9)$$

Clearly, the difference between HCMM and CMM turns out in different expressions for SA_{start} and SA_{dest} .

The derivation is simpler in the case of HCMM. First of all, it is easy to realise that the attractions of the starting and destination cells do not depend on the fact that the node is inside or outside the starting cell. The attraction to the starting (destination) cell depends only on the relationships with nodes having the starting (destination) cell as home, and on the number of such nodes. Thus, the attractions in HCMM are as follows:

$$\begin{cases} SA_{start}^{(H)} &= \frac{\sum_{j=1}^{n-1} w_{ij}}{fn} \simeq \bar{w} \\ SA_{dest}^{(H)} &= \frac{\sum_{j=1}^n p_r^{(n-1)} w_{ij}}{fn} \simeq \frac{p_r(n-1)\bar{w}}{fn} \end{cases} \quad (10)$$

Closed form expressions for the average time spent in the IN and OUT states in HCMM can be derived by replacing Equations 10 and 9 in Equation 8.

In the case of CMM computing attractions is more involved. The attraction to a cell dynamically depends on the number of nodes actually being in that cell. For the sake of simplicity, we carry on the analysis under the hypothesis that q nodes of the starting cell are roaming in the destination cell, and qt nodes of the destination cell are roaming in the starting cell. The attraction of the destination cell on a node currently roaming in the starting cell (and belonging to the starting cell's community) are computed based on the following line of reasoning. The node is attracted to the destination cell because nodes of its community are roaming there. Since links have been rewired, the node has links just towards a fraction of these nodes, i.e., towards $(1-p_r)q$ nodes, resulting in a contribution to the attraction equal to $\sum_{j=1}^{q(1-p_r)} w_{ij} \simeq q(1-p_r)\bar{w}$. The node is also attracted by nodes of the destination's community, to which it has been rewired. The probability of the node having been rewired to a random node of the destination community is $\frac{(n-1)p_r}{fn}$, and the number of nodes exerting such attraction is $\frac{(n-1)p_r}{fn}(fn-qt)$. Based on the above line of reasoning (applicable also to the attraction of the starting cell) it is possible to derive the required attractions formulas for CMM, as follows:

$$\begin{cases} SA_{start}^{(C,in)} &= \frac{(n-1-q)(1-p_r) + \frac{(n-1)p_r}{fn} qt}{n-q+qt} \cdot \bar{w} \\ SA_{dest}^{(C,in)} &= \frac{q(1-p_r) + \frac{(n-1)p_r}{fn}(fn-qt)}{fn-qt+q} \cdot \bar{w} \\ SA_{start}^{(C,out)} &= \frac{(n-q)(1-p_r) + \frac{(n-1)p_r}{fn} qt}{n-q+qt} \cdot \bar{w} \\ SA_{dest}^{(C,out)} &= \frac{(q-1)(1-p_r) + \frac{(n-1)p_r}{fn}(fn-qt)}{fn-qt+q} \cdot \bar{w} \end{cases} \quad (11)$$

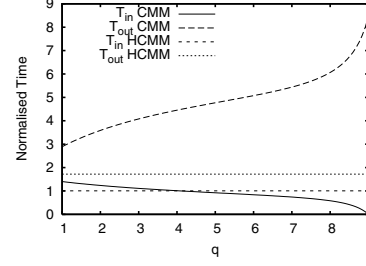


Fig. 13. Average time in the IN and OUT states as functions of q .

In the case of CMM the closed form expression of $E[T_{in}]$ and $E[T_{out}]$ is not as simple as in HCMM. The key point is the fact that in CMM these figures depend on the dynamic evolution of the users' movements. Specifically, they depend on q and qt , which are not model parameter, but change based on the nodes movements. Therefore, in CMM it is very hard to set model parameters to achieve a desired nodes' behaviour as far as nodes' physical positions. On the other hand, in HCMM $E[T_{in}]$ and $E[T_{out}]$ do *not* depend on the dynamic evolution of the system, but depend only on f, n, p_r , and p_e . This means that HCMM, while retaining the social theoretical approach of CMM, also provides simple knobs to control the time spent by nodes in the preferred physical locations. These remarks are confirmed by Figure 13, which plots $E[T_{in}]$ and $E[T_{out}]$ for CMM and HCMM as functions of q (time is normalised with respect to $\bar{T}^{(in)}$).

IV. CONCLUSIONS

In this paper we have provided a way to join two complementary and promising approaches to realistic mobility modelling for opportunistic networks: defining user movements based on i) human social relationships, and ii) attractions to users' preferred locations. The resulting Home-cell Community-based Mobility Model (HCMM) still shows well known statistical features observed in real traces, and provides simple knobs to control the time spent by users in their preferred locations.

REFERENCES

- [1] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic Networking: data forwarding in disconnected mobile ad hoc networks," *IEEE Communications Magazine*, vol. 44, no. 11, Nov. 2006.
- [2] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, 2002.
- [3] J. T. Camp and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communication and Mobile Computing*, vol. 2, no. 5, 2002.
- [4] M. Musolesi and C. Mascolo, "A Community Mobility Model for Ad Hoc Network Research," in *Proc. of ACM REALMAN*, 2006.
- [5] C. Tudeuce and T. Gross, "A mobility model based on WLAN traces and its validation," in *Proc. of IEEE Infocom*, 2005.
- [6] D. Lelescu, U. C. Kozat, R. Jain, and M. Balakrishnan, "Model T++: an empirical joint space-time registration model," in *Proc. of ACM MobiHoc*, 2006.
- [7] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," in *Proc. of Infocom*, 2006.
- [8] W. j. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling Time-Variant User Mobility in Wireless Mobile Networks," in *Proc. of IEEE Infocom*, 2007.
- [9] J. Ghosh, M. J. Beal, H. Q. Ngo, and C. Qiao, "On profiling mobility and predicting locations of wireless users," in *Proc. of ACM REALMAN*, 2006.
- [10] D. J. Watts, *Small Worlds The Dynamics of Networks between Order and Randomness*. Princeton University Press: Princeton Studies on Complexity, 1999.
- [11] A. Chaintreau, P. Hui, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. Mob. Comp.*, vol. 6, no. 6, pp. 606–620, 2007.
- [12] C. Boldrini, M. Conti, and A. Passarella, "Impact of Social Mobility on Routing Protocols for Opportunistic Networks," in *Proc. of IEEE WoWMoM AOC Workshop*, June 2007.