2011 Eighth IEEE International Conference on Mobile Ad-Hoc and Sensor Systems

On the Invariance of Spatial Node Density for Realistic Mobility Modeling

Bruno A. A. Nunes and Katia Obraczka
Department of Computer Engineering, Baskin School of Engineering
University of California, Santa Cruz, CA, USA
{*bnunes, katia*}*@soe.ucsc.edu*

***Abstract*—In this paper we show that human mobility exhibits "persistent" behavior in terms of the spatial density distribution of the mobile nodes over time. Using real mobility traces, we observe that the original non-homogeneous node spatial density distribution, where some regions may be quite dense while others may be completely deserted, is maintained at different instants of time. We also show that mobility models that select the next node position based on the position of other nodes, a la "preferential attachment", do not preserve the original spatial node density distribution and lead to behavior similar to random mobility as exemplified by the Random Waypoint model. To the best of our knowledge, this is the first time that these phenomena have been reported. Based on these observations, we propose a simple mobility model that preserves the desired spatial density distribution. Moreover, when simulating the operation of a network moving according to the proposed model, we found that performance results expressed by a number of network metrics also match closely results obtained under mobility governed by real traces. We also compare our results to models whose steady-state do not preserve the original non-homogeneous density distribution and show that network performance under such regimes deviates from performance under real trace mobility.**

I. Introduction

Mobility models are an indispensable tool in the design, testing, and evaluation of wireless networks and their protocols. This is the case for infrastructure-based networks (e.g., wireless LANs or WLANs), but even more so in infrastructure-less networks, a.k.a., wireless, self-organizing networks (WSONs) which include wireless mobile ad-hoc networks (MANETs), wireless sensor networks (WSNs) and disruption-tolerant networks (DTNs). In these networks, unlike their infrastructure-based counterparts where only the end user node is mobile, every node may have unlimited mobility. As an indication of the importance of mobility models to the study of wireless network protocols, most well-known network simulators include "mobility generators", which, following a pre-specified mobility regime, determine the position of network nodes over time during simulation runs. Synthetic mobility generators have been extensively used in the study of wireless networks [1]. A notable example is the Random-Waypoint Mobility (RWP) regime.

Recently, however, network researchers and practitioners have been trying to use more realistic scenarios to drive the evaluation of wireless network protocols. This motivated initiatives such as the CRAWDAD [2] trace repository, which makes real traces available to the networking community. These traces can then be used to run trace-driven simulations. Even though initiatives like CRAWDAD have greatly increased availability of real traces, the number and variety of publicly-available mobility traces are still quite limited. Consequently, relying exclusively on traces to design and evaluate network protocols would not allow a broad enough exploration of the design space.

To address this problem, a number of efforts have proposed mobility models based on realistic mobility patterns [3]. Notable examples include [4, 5, 6, 7]. Alternate approaches to developing realistic mobility generators try to use characteristics of real human mobility. For example, some efforts employ "transition probabilities", while others are based on maps of the area being simulated [8].

More recent work focuses on the "scale-free" properties observed in many real networks like the Internet, the Web, and some social networks, to name a few. The seminal work of Barabási and Albert [9] proposes a model that generates scale-free networks, i.e., networks whose node degrees follow a power law distribution. One key concept underpinning the Barabási-Albert model is referred to as the *preferential attachment* principle which states that "the more connected a node is, the more likely it is to receive new links". As described in Section VI, which discusses related work in more detail, several recently proposed mobility models (e.g., [10, 11, 12, 13, 14, 15]), try to mimic real human mobility by following the preferential attachment principle: they define *attraction points*, whose probabilities of attracting other nodes increase as more nodes congregate around them.

In this paper, we show that using the preferential attachment principle for modeling human mobility leads to undesirable long-term behavior, which does not preserve

978-0-7695-4469-4/11 $26.00 © 2011 IEEE
DOI 10.1109/MASS.2011.39

the original spatial node density distribution and leads to behavior similar to random mobility as exemplified by the Random Waypoint model. Instead, real human mobility exhibits "persistent" density heterogeneity as exemplified in Figure 1. This figure shows the spatial density distribution for one of the traces used in this paper which was collected in the Quinta da Boa Vista Park in Rio de Janeiro, Brazil. This distribution is defined as the percentage of cells containing $\geq k$ nodes. The 5 curves in the plot refer to the initial distribution and distributions at instants 300, 500, 700, and 900 seconds, which is the end of the trace collection interval. From the graph we observe that the density distribution does not vary much with time: the largest deviation from the initial distribution for any value of k at any instant is 8.3%; the average deviation from the initial distribution measured in all the instants for all values of k is 1.27%. Similar observations can be drawn from the other traces used in our work as reported in Section III-A. We should point out that, from a survey we conducted on 82 papers published in various networking conferences between 1996 and 2010 which used some form of simulation study, only 57 papers listed simulation duration time. Out of those, 21 papers used 900s, 12 papers used 500s, 4 used 600s, and 3 used 300s. All others reported different times all lower than 300s. This means that the time windows that we use in this paper to study human mobility are representative of– and go beyond the simulation times used to study wireless networks and their protocols.

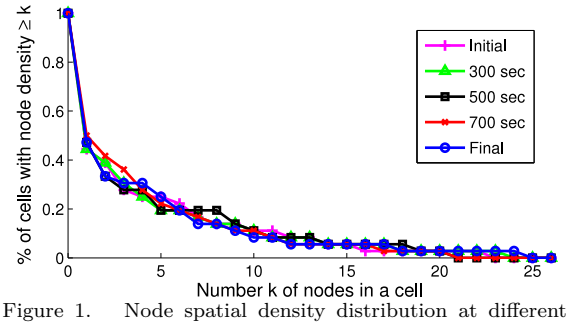

Figure 1. Node spatial density distribution at different trace collection times for mobility in a city park.

Based on our observations of the persistent density distribution behavior in human mobility, we then propose a simple, yet realistic model of human mobility called *Intensity Waypoint*, or *IWP*. Besides preserving, by design, the invariance of the node density distribution over time, IWP is simple and easy to use and implement. We argue that these latter two features of IWP are critical for its widespread use and adoption, vis-a-vis the RWP model. Another important feature of IWP is its generality, i.e., it can be used to model mobility in any kind of networking environment including Wireless LANs or ad hoc networks.

We evaluate IWP in two different ways: first, we investigate its ability to adhere to the node density distribution invariance property. To this end, we define the *node density distribution* metric (mentioned above) and use it to compare IWP against RWP and the Natural Mobility Model [12], a representative of mobility models based on preferential attachment. In this study, we use three GPS-based mobility traces collected in different scenarios, namely at a park, a University campus, and a state fair. The traces were used to set the parameters of the models as well as the baseline for the comparative study. The second part of the evaluation focuses on studying the performance of core network functions, in this case routing, under IWP mobility. The same traces described above were used as baseline to compare IWP against RWP and the Natural Mobility Model. Our results show that, by preserving the original node density heterogeneity, IWP is able to closely approximate real network behavior according to a number of routing performance metrics.

In summary, to the best of our knowledge, this is the first time that the invariance of the non-homogeneous spatial density distribution of human mobility has been observed. This is also the first work to point out that mobility models based on the preferential attachment principle does not preserve the original spatial node density distribution and lead to behavior similar to random mobility, a la Random Waypoint model.

The remainder of this paper is organized as follows. Our model IWP is presented in detail in the next section, and Section III describes how its parameters are extracted from real traces. Our experimental methodology for evaluating IWP is described in Section IV, and our results presented in Section V. State of the art is presented in Section VI, while we discuss the impact of our work given previous work in the field. Finally, Section VII concludes the paper.

II. Intensity Waypoint Mobility

The main design principles guiding the proposed model are: (1) preserving the invariance of the original node density distribution and consequently its inherent non-homogeneity, (2) generality, i.e., applicability to different types of network scenarios, and (3) simplicity and ease of use in order to facilitate the model's adoption.

In our model, we divide the simulation area in equal sized squares, that we refer to as cells. Each cell $i \in \{1..N\}$, where N is the total number of cells composing the simulation area, is assigned an intensity $\mu_{i,t}$ at time t given by an *Intensity Map (IM)*. The IM is a vector composed of N elements, where each element has a value $\{\mu_{i,t} \in \mathbb{R} \mid \mu_{i,t} \geq 0\}$ that indicates how intense the activity in a cell is. We define the IM and how we build it in Section III-B.

The probability $p_{i,t}$ that cell i is chosen as the next waypoint at time t is given by Equation 1 below:

$$p_{i,t} = \frac{\mu_{i,t}}{\sum_{j=1}^{N} \mu_{j,t}}, \forall i \in \{1..N\} \quad (1)$$

The evolution of the spatial node density distribution over time is an important aspect of the network, but as our results show, this distribution does not vary much with time. For that reason, we make $\mu_{i,t}$ and $p_{i,t}$ time-independent; thus, $p_{i,t} = p_{i,t'} = p_i$.

We define a step function as in Equation 2. This function will allow us to set up the attraction areas by simply setting the values of x_0 to the physical limits of a given attraction area or *cell i*: $x_{i,min}$, $x_{i,max}$, $y_{i,min}$ and $y_{i,max}$. The cell's area A_i is given by $(x_{i,max} - x_{i,min}) \times (y_{i,max} - y_{i,min})$.

$$u(x - x_0) = \begin{cases} 1 & \text{if } x \geq x_0 \\ 0 & \text{if } x < x_0 \end{cases} \quad (2)$$

Thus, the probability distribution of selecting the next destination (x_d, y_d) is defined by Equation 3 as follows:

$$f(x_d,y_d) = \frac{1}{\sum_{i=1}^{N} p_i A_i} \Bigg[\sum_{i=1}^{N} p_i \Bigg[\Big(u(x_d - x_{i,min}) - u(x_d - x_{i,max}) \Big) \Big(u(y_d - y_{i,min}) - u(y_d - y_{i,max}) \Big) \Bigg] \Bigg] \quad (3)$$

Equation 3 can be viewed as a general form of the one given in [27] for a single attraction point. According to this equation, in order to select a node's new destination, we need a mechanism that allows us to differentiate the choice of the next way-point based on the intensity of each cell. We turn to a *Fitness Proportionate Selection*-based scheme known as *roulette-wheel selection* for that. This strategy is commonly used in genetic algorithms [16].

Based on the values of p_i, a circular wheel is divided into N sectors, like a pie chart. The center angle of the i-th sector is $2\pi p_i$, which makes the area of each sector proportional to the probability p_i. The roulette wheel selection strategy can then be described in two steps, every time the next cell is to be selected. First, generate a random number n uniformly distributed, where $0 \leq n \leq 1$ and second, if $p_1 + p_2 + \ldots + p_i \leq n < p_1 + p_2 + \ldots + p_{i+1}$, then cell i is selected to be the next one visited by the node. This is equivalent to spinning the roulette and having it stop at sector i. The bigger p_i is, the bigger the sector and the bigger are the chances of sector i being chosen. Once the new cell i is known a coordinate (x, y) is then uniformly chosen inside of cell i, and $x_{i,min} \leq x \leq x_{i,max}$ and $y_{i,min} \leq y \leq y_{i,max}$. Hence, the cumulative probability q_i of choosing a cell is given by Equation 4.

$$q_i = \sum_{j=1}^{i} p_j, \forall i \in \{1..N\} \quad (4)$$

The same mechanism described above can also be adopted for choosing nodes' velocity and pause time. In this case, instead of an IM we would use a velocity domain vector $V = \{\nu_1, .., \nu_{n_v}, .., \nu_{N_v}\}$ containing all the N_v instances ν_{n_v} of nodes' velocity measured in the traces and a velocity intensity vector $F_v = \{f_1, .., f_{n_v}, .., f_{N_v}\}$ containing the frequency f_{n_v} of all velocity instances ν_{n_v}. Analogously, we would also have a pause time domain vector $P = \{t_1, .., t_{n_p}, .., t_{N_p}\}$ containing all the N_p instances t_{n_p} of nodes' pause time and a pause time intensity vector $F_p = \{f_1, .., f_{n_p}, .., f_{N_p}\}$ containing the frequency f_{n_p} of all pause time instances t_{n_p}. The elements of V and P would identify the sectors of the roulette-wheel, where velocity and pause time wheels would have N_v and N_p sectors respectively, and the elements of F_v and F_p would determine the size of each sector. However, in this paper we wanted to investigate the formation of clusters and the non-uniform nature of node density in real scenarios and its impacts on networking. Thus, we focus on the intensity of attraction areas and use uniform distributions to model node velocity and pause time.

III. Building the Intensity Map

Acknowledging the challenge in defining what are *real* mobility characteristics, we approach this subject by extrapolating reality from the observation of real human mobility. To that end, we extract node density characteristics form the traces and represent it in our model in the Intensity Map, or IM. The IM can be viewed as a "density map", i.e., a map of the region over which nodes move showing the node density at each cell. In this section, we describe how to build the IM from real traces. Note that once an initial IM is built for a specific mobility scenario, variants of that scenario can be generated by controlling the density distribution reported in the trace-based IM. For example, if one wants to further intensify certain cells or increase/decrease the number of "hotspots", that can be easily accomplished by adjusting the IM accordingly.

In this paper, we use three GPS traces collected in scenarios that are quite diverse, namely a park, a University campus, and a fair. These traces are presented in detail below. Then, we describe how we extracted density information from the traces and use this information to drive IWP mobility.

A. Real Traces

Table I summarizes the GPS traces in terms of number of users, duration of the trace, and the GPS sampling period.

Trace	# users	Duration	Samples
Quinta [17]	97	900s	1s
KAIST [18]	78	5000s	10s
Statefair [18]	19	8000s	10s

Table I
SUMMARY OF THE GPS TRACES STUDIED.

Quinta, refers to the "Quinta da Boa Vista Park" trace, first presented in [17]. It is a GPS trace collected at a park in the city of Rio de Janeiro, Brazil. The park has many trees, lakes, caves, and trails. It houses the National Museum of Natural History and the city Zoo. The *KAIST* trace [18], on the other hand, is a GPS trace collected at a University campus environment (KAIST) in Daejeon, South Korea. The *Statefair* trace, also available at [18], is yet another mobility scenario showing daily GPS track logs collected from the NC State Fair held in North Carolina, USA.

Given that the traces of some of the nodes presented some discontinuity, we selected sections of the raw traces where no discontinuity occurred. We selected only nodes where we could find a continuous sequence of GPS fixes that were as long as 900, 5000, and 8000 seconds for the Quinta, KAIST, and Statefair traces, respectively. These were the total durations of the traces.

B. Extracting Information from Traces

We extract from the traces the distributions of speed, pause time, and node density by using the trace's sampling period. For example, in the Quinta trace, the sampling period is $T = 1$ seconds. A histogram was built with this information, for speed and pause time, giving us the relative frequency at which each value occurs. We define the node's speed as $\frac{d}{\Delta t}$ where d is the distance traveled between two consecutive entries in the GPS trace at times t_1 and t_2 and $\Delta t = t_2 - t_1$. Pause time is defined as $P = \Delta t$, if $d < threshold$, or zero otherwise. The threshold is not zero here to account for GPS error. We set this threshold to be 2 meters for KAIST and Statefair traces and 0.5 meter for the Quinta trace, due to jitter in GPS update frequency.

To extract spatial node density, the area is divided into squared cells of 50 x 50 meters. Using this particular cell size to generate the density map was decided empirically in our experiments: it provided both adequate resolution as well as node clustering. An alternative approach could be identifying specific attraction zones, as was done in [11]. This is one of the topics of future work we plan to address. For this first version of IWP, a fixed size cell was used.

With the area divided into cells, we took a snapshot of the nodes' spatial distribution in the area being simulated every T seconds. The value of $T = 10$ was used since, for the size of the cells and the speeds sampled from the traces, a node could not on average change between more than two cells during T. For every cell, at every interval T we counted the number of nodes in each cell. We then averaged the number of nodes in each cell over the course of the whole duration of the trace, giving us what we will call here, a *Intensity Map (IM)*. The IM gives the popularity of each cell, i.e., how intense is the node mobility activity in a cell, or in other words, how interested nodes are in general in each cell.

The cumulative probability of choosing a cell, given by Equation 4 can be visualized in Figure 2 for Quinta and Statefair scenarios and for RWP (uniform) and IWP (set by the IM extracted from the traces). The curves for the KAIST scenario are similar, but were omitted for the sake of clarity of the plot.

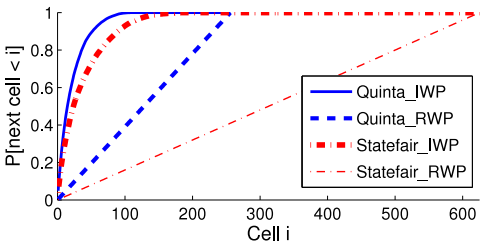

Figure 2. Cumulative cell selection distribution.

We acknowledge that always relying on extracting the IM from available traces can be somewhat constraining since there is no large and diverse enough set of publicly available traces. In general, the ones that exist represent a limited set of scenarios such as University or company campuses, parks, vehicles in city streets, conferences, etc. An alternative to using real traces to extract the IM is to use a technique like the one introduced in [19]. In this work, the authors propose a way to generate non-homogeneous spatial node distributions. Using such a technique, one could generate an initial non-homogeneous placement of nodes, which then could be used to generate the IM to feed our model by averaging the number of nodes in each cell. The fact that this technique does not allow averaging density over time does not pose as a big drawback: as we will show in Section V, spatial node density distribution does not vary much in time.

Our model not only manages to preserve a desired level of heterogeneity over time [20], but allows for tunning the desired spatial density for evaluating network protocols under different levels of node density, according to the researchers needs. In case the objective is to simulate some specific real scenario, information on trace sources and repositories can be found in [21] and

can be used to generate the IM as described here.

IV. Evaluation Methodology

In this section, we describe how we investigate IWP's behavior and evaluate its performance. The baselines used for our comparative study are the random waypoint mobility model, or RWP, and the natural mobility model. We then start by describing these two mobility models.

A. RWP Mobility

In the RWP model, mobile nodes are initially placed in the area being simulated according to a given distribution. Typically, a uniform distribution is used. Each node remains in its position for a given period of time, called *pause time* P, uniformly chosen in the interval $[0, P_{max}]$, where P_{max} is a parameter whose value needs to be pre-specified. After this period, the mobile node chooses a new destination (x_d, y_d) uniformly distributed in the simulation area, and a velocity, also uniformly distributed in the interval $[v_{min}, v_{max}]$, where both v_{min} and v_{max} are pre-specified parameters. Once the destination is reached, the node pauses again and chooses another destination and velocity, as described above.

B. Natural Mobility Model

We also compare IWP's behavior and performance against a model that follows the preferential attachment principle. As representative of this family of models we use the natural mobility model, or simply *Natural* [12].

As discussed in Section VI, Natural is based on attraction points, where the attractiveness of each point is proportional to its popularity (i.e., number of nodes at or going towards it) and inversely proportional to the distance to it. Thus, the probability $\Pi(a_i)$ that an node z_k chooses an attractor a_i among all possible attractors is proportional to the portion of the total attractiveness it carries: $\Pi(a_i) = \frac{\mathcal{A}_{a_i,z_k}}{\sum_j \mathcal{A}_{a_j,z_k}}$. The attractiveness of an attractor is then defined as:

$$\mathcal{A}_{a_i,z_k} = \frac{(1 + \sum_{z_j \in \mathbb{Z}, z_j \neq z_k} B(a_i, z_k))}{\sqrt{(X_{a_i} - X_{z_k})^2 + (Y_{a_i} - Y_{z_k})^2}} \qquad (5)$$

where $B(a_i, z_k)$ is a Bernoulli variable, with $B = 1$ if the individual z_k is going toward or staying at attractor a_i and 0 otherwise, and X and Y are the coordinates of a node and an attractor. In our implementation, we divided the simulation area in equally sized squares, or cells, and consider each cell to be an attraction point. The coordinates (X_{a_i}, Y_{a_i}) mark the center of the i-th attraction point. Once the new destination is known, the node travels towards it with a velocity uniformly distributed in the interval $[v_{min}, v_{max}]$. A pause time is randomly selected once arriving at the destination before choosing another destination and beginning the process again.

C. Simulation Setup

We ran two types of simulations: first, we used a modified version of the Scengen [22] scenario simulator to generate traces according to RWP–, Natural–, and IWP mobility. With those traces we generate the node density distribution results presented in Section V-B. In the second type of simulation, once the synthetic traces were generated, these and the GPS traces were fed to the Qualnet network simulator [23] in order to evaluate their impact on core network functions, like routing.

In the first type of experiment, three sets of synthetic traces were generated using the RWP, Natural and IWP mobility models. The velocity range was set in a way that the average velocities would match the ones measured in the GPS traces for Quinta, KAIST and Statefair. In order to address the decaying velocity problem reported in [24], we followed the recommendations mentioned in that work. To that end, the velocity range was set to be $\pm$ the standard deviation measured in the real traces around the measured average velocity. Thus, the velocities were chosen *uniformly* in a range in which the lower limit was greater than zero and where the mean matches the one measured in the real traces. This is not an optimum solution and it is the simplest solution mentioned in [24], but since our focus is not on evaluating the routing protocol itself and we rather want to compare the mobility models, we found that solution suitable for these purposes.

Pause time was chosen uniformly in the range $[0, P_{max}]$, where the value of P_{max} was set to an appropriate value, in a way that the average pause time would match the one measured in the real traces. The same was done for the dimensions of the rectangular simulation area, set to be the same as in the GPS traces. Moreover, in all simulation scenarios, Quinta, KAIST and Statefair, we used the same initial positions found in their respective real traces for the same number of users. For further discussions on the actual distributions for these traces' mobility parameters, please refer to [17, 18].

In the RWP simulations, the next destination (x_d, y_d) was chosen uniformly over the simulated area. For the IWP simulations, the choice of (x_d, y_d) was given by Equation 3, where the intensity values μ were measured from the real traces as described in Section III-B. For Natural, the probability of choosing a given cell is computed "on-the-fly", based on the cell's attractiveness.

We of course do not attempt to simulate the real scenarios exactly with the RWP as it is. It was an attempt to approach the simulated mobility patterns to the real scenarios so that we could compare the simulation results driven by randomly generated mobility traces against

simulations over a real trace. The goal is to evaluate how far apart the results using the RWP are form the ones using the real trace and how close we can get by using the proposed model. Table II summarizes the simulation parameters.

Parameter	Quinta	KAIST	Statefair
Avg. Speed ($\pm\sigma$)(m/s)	1.2($\pm$0.53)	0.72($\pm$0.68)	0.48($\pm$0.39)
Avg. Pause (sec)	3.6	86	72
Area (meters x meters)	840x840	5000x5000	1260x1260
Duration (sec)	900	5000	8000
# nodes	97	78	19
# CBR flows	20	20	10

Table II
SIMULATION PARAMETERS.

The traffic scenarios used in our second type of simulations are composed by *20 CBR flows* with randomly chosen source and destination pairs, for the Quinta and KAIST scenarios. *10 CBR flows* were used in the Statefair simulations. Flows start at randomly chosen times and stay active during the course of the whole simulation, at a *rate of 4 packets per second*. The routing protocol used to compare the network metrics between the different mobility patterns is the Ad-Hoc on-demand distance vector (AODV) protocol. The *radio range* was set to 150m, for 802.11b data rate of 11.0 Mbps.

Results are reported here with a *confidence interval* of 90% and reflect *10 runs* of the simulations using the real GPS traces. Since we could not change the mobility for the simulations with the GPS traces, we varied the traffic scenarios by changing the source and destination pairs of the flows in each of the 10 runs. The same traffic scenarios were used to feed the RWP, Natural, and IWP simulations, but in these cases, we generated 5 mobility traces with each model, giving a total of $10 \times 5 = 50$ simulation runs for each one of the synthetic models.

V. RESULTS

In this section we define our metrics and present our results for node spatial density and network routing.

A. Spatial Node Density

In order to study spatial node density behavior, we define the following metric:

- *Node density distribution:* the percentage of cells containing $\geq k$ nodes.

Figures 3, 4 and 5 show the density distribution for one of the traces and each mobility model, namely IWP, RWP, and Natural. The curve labeled "*Initial*" shows the trace's density distribution at the beginning of the trace collection interval, while the curve labeled with the trace's name shows the distribution at the end of the trace collection interval, namely 900, 5000, and 8000 seconds for the Quinta, KAIST, and Statefair traces, respectively.

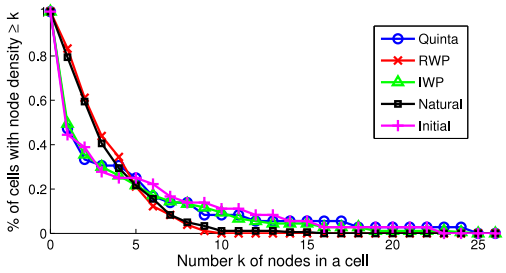

Figure 3. Node density distribution for the Quinta trace.

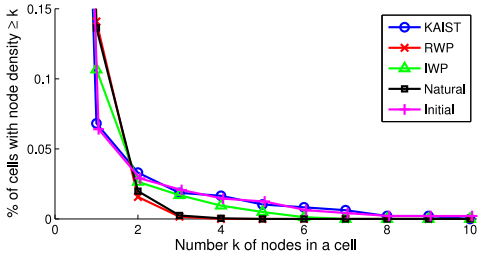

Figure 4. Node density distribution for the KAIST trace.

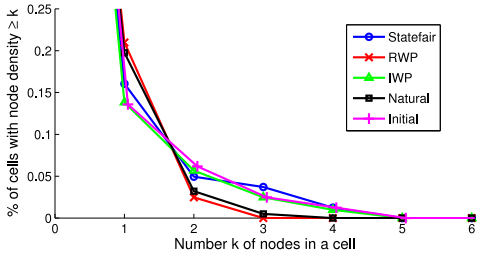

Figure 5. Node density distribution for the Statefair trace.

The data points shown for the RWP, Natural, and IWP models were averaged over 10 runs. The plots for the $KAIST$ and $Statefair$ traces are zoomed in to the region of interest. In those two plots, the only point not shown is $k = 0$, where the percentage of cells containing 0 or more nodes $P[k \geq 0]$ is the same for every curve and it is, of course, equal to 100%.

From these plots we observe that IWP's density distribution follows closely the distribution of the real trace, for all three traces. In the case of RWP, the majority of cells present similar number of nodes, and no cells contain significantly greater concentration of nodes. This is also the case for Natural. In the beginning of the simulation the initial distribution weighs on the computation of the attractiveness of the cells, but as the simulation goes on, more and more nodes may choose with small, but greater than zero probability, cells where there are no other nodes present. Over time, this behavior spreads the nodes around different cells, leveling the attractiveness of cells in general, approaching uniformity and RWP's density distribution.

An important result here is to see how in real traces the spatial density distribution does not change much

over time preserving the original non-homogeneous density distribution. Moreover, using IWP, we are able to preserve the initial spatial distribution and its non-homogeneity and thus yield very similar density distribution behavior when compared to the real trace. That holds true for all the three trace scenarios studied.

B. Network Routing

The metrics used in our comparative study of network routing performance under the IWP mobility regime are:

- *Throughput*: Total number of bits received at the destination divided by the time elapsed between the reception of the first data packet and the reception of the last data packet. This quantity is then averaged for all nodes.
- *End-to-End Delay*: Time elapsed between data transmission and reception then averaged for all nodes.
- *Delivery Ratio*: Ratio between the total number of packets received by all nodes and the total number of packets sent by all nodes.
- *Routing Overhead*: Total number of routing control packets transmitted by all nodes averaged over nodes.

Figures 6, 7, 8, and 9 show, respectively, throughput, delay, delivery ratio, and routing overhead over time for the Quinta scenario. Notice the clear discrepancy between the real trace results and the RWP model. Also notice how the discrepancy widens over time. Routing performance suffers under RWP mobility once it spreads the nodes randomly over the simulation area. IWP, on the other hand, like the real trace, allows the formation and preservation of clusters, which, in the case of this scenario, ensures close to 100% delivery ratio.

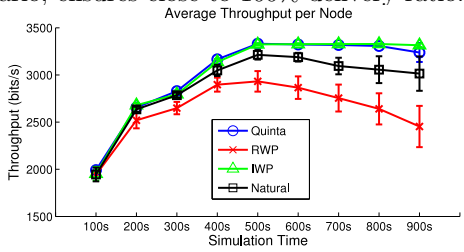

Figure 6. Throughput for the Quinta Scenario.

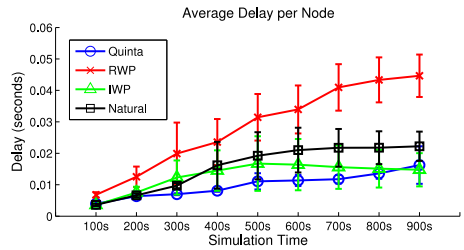

Figure 7. Delay for the Quinta Scenario.

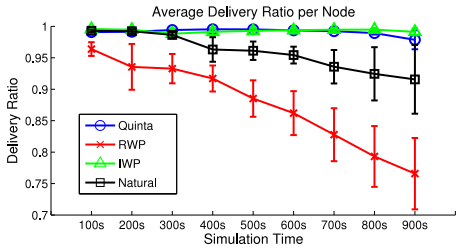

Figure 8. Delivery Ratio for the Quinta Scenario.

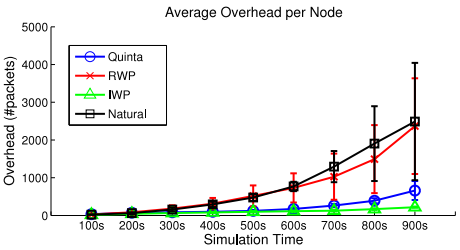

Figure 9. Routing Overhead for the Quinta Scenario.

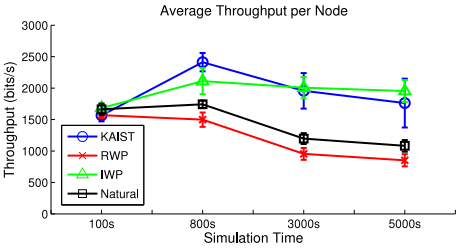

Figure 10. Throughput for the KAIST Scenario.

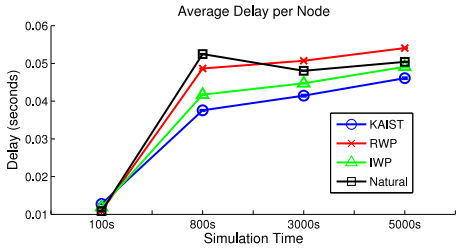

Figure 11. Delay for the KAIST Scenario.

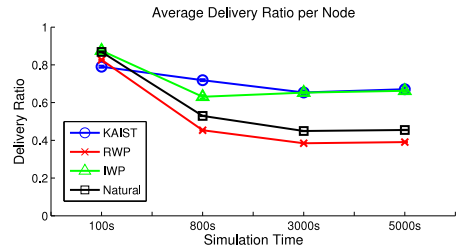

Figure 12. Delivery Ratio for the KAIST Scenario.

In the case of Natural mobility model, specifically for the Quinta scenario, we notice that routing performance for this model stays close to the real trace curves up until around 300 seconds. Before that time, the probabilities of choosing each cell are differentiated by the initial non-uniform node positions, and the model is capable

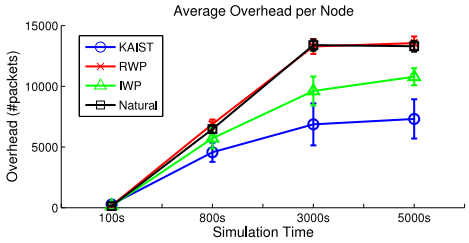

Figure 13. Routing Overhead for the KAIST Scenario.

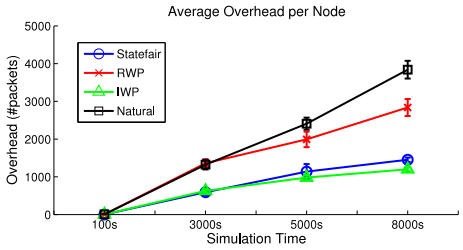

Figure 17. Routing Overhead for the Statefair Scenario.

of maintaining some spatial node aggregation. However, after this point, nodes that were starting to roam away choose new cells where node concentration is not as high. As more and more nodes start to spread throughout the simulation area, the probability of choosing a new cell starts approaching a uniform distribution. This behavior causes the clusters to dissipate and the routing performance starts to diverge from the real traces.

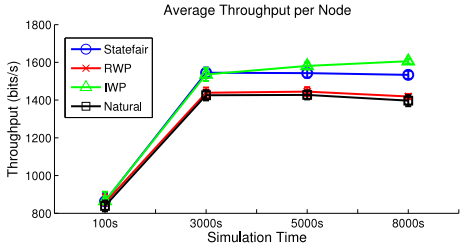

Figure 14. Throughput for the Statefair Scenario.

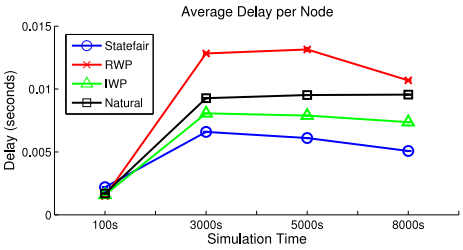

Figure 15. Delay for the Statefair Scenario.

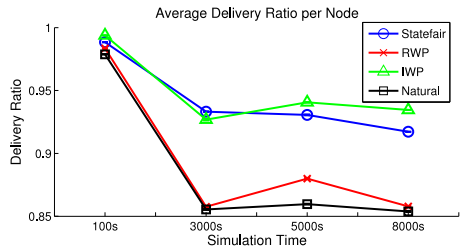

Figure 16. Delivery Ratio for the Statefair Scenario.

Figures from 10 to 13 and from 14 to 17 show routing performance metrics for the KAIST and Statefair scenarios, respectively. Once again, as in the Quinta scenario, we can see the IWP curves diverging from the RWP curves and approaching the real trace curves. In all the scenarios the routing overhead for RWP and Natural is much greater than for IWP and the real traces. This is mainly due to paths breaking caused by the random movement of RWP nodes and the increasing randomness towards the end of the simulations for Natural. Also, delivery ratio for these two scenarios is lower than for the Quinta scenario given the lower relative density caused by a larger simulation area and increased average distance between nodes. Nonetheless, the negative impact of a more sparse network in the case of KAIST and Statefair scenarios is greater for RWP and Natural. For these two scenarios, we were not able to see similar behavior to the real traces in the beginning of the simulations for Natural, as we saw in the Quinta scenario, due to the larger and lower resolution time scale. Nonetheless, the way Natural diverges from the real traces towards the end of the simulations remains.

Approaches like [10, 11, 12, 13, 14, 15], where the probability of choosing the new destination changes over time depending on nodes' positions, can present long transient phases and a stead state similar to a completely random behavior such as observed when applying the RWP model. We argue here that being able to set the spatial distribution in the beginning of the simulation and being able to control it is important, and the proposed approach allows it.

VI. Related Work

There has been considerable work on trying to depart from random mobility models towards more realistic ones applied to the evaluation of wireless networks and their protocols. In this section, we do not try to present a complete survey of the state-of-the-art on realistic mobility modeling; rather, our goal is to highlight the efforts that are more closely related to our work and thus motivated it.

Some existing approaches use RWP mobility as a starting point, especially due to its simplicity and ease of use and implementation. Another way to differentiate existing approaches is in terms of the type of network environment they target, namely infrastructure-based and infrastructure-less networks. An example of mobility model for infrastructure-based network, more

specifically WLANs, is the work described in [25] which captures mobility as nodes move between access points. Parameters for the model are extracted from real WLAN traces. In [11], Kalman Filters (KF) are used to extract paths taken by nodes as they move from a hotspot to another in the Dartmouth campus. The waypoints in the path between hotspots are given by the KF for the map of that specific scenario. Hotspot regions are identified using a Gaussian distribution at each pause location, weighted by the time a node spends on each hotspot and then averaged over all nodes. Our mobility model, IWP, tries to yield high fidelity to real mobility and can be applied to both types of networks.

As mentioned above, several efforts have focused on proposing variants of RWP mobility. For instance, to avoid RWP's "zig-zag" behavior, in [26], a model which avoids sudden accelerations and sharp turns during direction changes has been proposed. Spatial node distribution of the RWP model has been studied in [27], where analytical expressions of the expected node distribution in one dimension have been derived as well as an approximation for the two-dimensional case. Another study of node density has been presented in [28], where real traces of vehicular mobility in a city have been used to find correlation between node density and speed. A model called "heterogeneous random walk" has been proposed, and node clustering under such mobility regime has been studied.

The scale-free properties of real networks were investigated in several studies. As pointed out in Section I, an important milestone on the subject is the work by Barabási and Albert [9] where it has been demonstrated that many real large networks are scale free, that is, the node degree in the network graph follows a power law. The authors discuss the mechanism responsible for the emergence of scale-free networks and argue that understanding this problem will require a shift from modeling network topology to modeling "network assembly and evolution". To this end, they define the Barabási-Albert model based on *growth* and *preferential attachment*. Growth refers to the fact that the number of nodes in the network increases over time, where a new node is placed with m edges connecting it to other m nodes. Preferential attachment means that a node will choose to connect to another node i with probability $\Pi(k_i) = \frac{k_i}{\sum_j k_j}$ based on the degree k_i of node i and any node j connected to node i.

Several mobility models have been inspired by the Barabási-Albert preferential attachment principle. For example, in [12], a model based on preferential attachment has been proposed, where the choice of going towards an attraction region is weighted proportionally to the region's popularity (i.e., the number of other nodes that chose it) and inversely proportional to the distance to it. The work proposed in [15] is another example of a model that follows Barabási-Albert's growth and preferential attachment principles. The authors even show a figure where they present their initial (after growth) and steady-state spatial distribution. It is possible to see how clusters dissipate and fade away over time. The same concept is also used in [13] where nodes are also driven by identified social interactions. Authors validate their approach by means of showing the power-law exponential decay of inter-contact time among the formed communities and compare it with measurements in real traces.

Bettstetter et al. point out that random mobility leads to homogeneous node distributions. In [19], they propose a method that creates initial non-homogeneous node distributions and in [20], a variant of RWP mobility that maintains the non-homogeneity of an original node distribution.

Map based approaches are interesting in the sense that they limit a node's roaming area, which is what happens in reality (i.e., vehicles in roads, people walking in trails in a park, campus streets and corridors in a conference or fair). For example, in [8], transitions between different locations in a campus map are modeled. The map is represented by a graph and users move between nodes of the graph, representing access points, buildings, and intersections in the map.

More recent approaches focus on the aspects of social interaction between nodes. Other examples of mobility modeling for social networks include the work in [29] and [14], where node movement is influenced by the strength of social ties and the choice of an attraction point is based on the history of visits of other nodes to that location.

VII. Conclusions

We investigated in this paper the spatial density distribution properties of real mobility using mobility traces collected in a diverse set of scenarios. More specifically, we chose three scenarios where human mobility was registered using GPS tracking devices carried by users as they move around a city park, a University campus, and a State Fair event. By defining a metric called spatial density distribution, we showed that, for all three scenarios, the original non-homogeneous node density distribution does not change over time. We also show that mobility models that rely on computing the probability of choosing a node's new destination dependent on the position of the other nodes can exhibit long transient phases and a steady state similar to completely random behavior as observed when applying the RWP model. To the best of our knowledge, this is the first time that these observations have been reported.

Acknowledging the importance of preserving and being able to control node density over time, we proposed a mobility model called Intensity Waypoint (IWP) that takes a density map as a parameter and where the mobility description is as simple as in the RWP model. By feeding the IWP model with a density map extracted from mobility traces, we were able to match closely the density distribution generated by the IWP to the one generated by the real traces, also preserving the original node density distribution. Moreover, when simulating the operation of a network moving according to the proposed model, we found that performance results expressed by a number of network metrics also match closely results obtained under mobility governed by the real traces. We also compared our results to models whose steady-state tend to random mobility and show that network performance under such regimes deviates from performance under real trace mobility.

Other features of our model are its generality as well as the user's ability to control node density distribution in order to simulate specific scenarios. Our conjecture is that, given the simplicity and ease of use of our approach, it has the potential of becoming adopted and make its way to be part of well-known network simulation platforms.

As directions of future work, we plan to investigate human mobility behavior over extended periods of time as well as proving mathematically what we observed empirically, i.e., the steady-state behavior of preferential attachment based mobility models as well as IWP's.

Acknowledgements

Financial support was granted by the CAPES Foundation Ministry of Education of Brazil, Caixa Postal 250, Brasilia - DF 70040-020 BrazilThis work was also partially supported by NSF grant CCF-091694 and a US Army-ARO MURI grant. Authors would like also to thank Bruno Sanso, Athanasios Kottas and Kumar Viswanath for their invaluable contribution to this work.

References

[1] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, pp. 483–502, 2002.

[2] CRAWDAD, "http://crawdad.cs.dartmouth.edu/."

[3] M. Musolesi and C. Mascolo, "Mobility models for systems evaluation," in *Middleware for Network Eccentric and Mobile Applications*, pp. 43–62, Springer, 2009.

[4] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan, "Characterizing user behavior and network performance in a public wireless lan," in *ACM SIGMETRICS 2002.*

[5] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," in *MobiSys 2003*, pp. 303–316.

[6] D. Tang and M. Baker, "Analysis of a local-area wireless network," in *ACM MOBICOM 2000.*

[7] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *ACM MOBICOM 2004.*

[8] J. Yoon, B. D. Noble, M. Liu, and M. Kim, "Building realistic mobility models from coarse-grained traces," in *MobiSys 2006.*

[9] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, 2002.

[10] A. Jardosh, E. Belding-Royer, K. Almeroth, and S. Suri, "Real-world environment models for mobile network evaluation," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 3, pp. 622 – 632, 2005.

[11] M. Kim and D. Kotz, "Extracting a mobility model from real user traces," in *IEEE INFOCOM*, 2006.

[12] V. Borrel, M. D. de Amorim, and S. Fdida, "On natural mobility models," in *WAC*, 2005.

[13] A. Mei and J. Stefa, "Swim: A simple model to generate small mobile worlds," in *IEEE INFOCOM 2009.*

[14] M. Musolesi and C. Mascolo, "Designing mobility models based on social network theory," *ACM SIGMOBILE Mobile Computing and Communication Review*, vol. 11, pp. 59–70, 2007.

[15] S. Lim, C. Yu, and C. Das, "Clustered mobility model for scale-free wireless networks," in *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*, 2006.

[16] Z. Pan, K. L.S., and Y. Chen, *Evolutionary Computation.* 2000.

[17] T. Azevedo, R. Bezerra, C. Campos, and L. de Moraes, "An analysis of human mobility using real traces," in *IEEE WCNC 2009.*

[18] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Kim, "CRAWDAD http://crawdad.cs.dartmouth.edu/ncsu/mobilitymodels/GPS/KAIST v.2009-07-23," July 2009.

[19] C. Bettstetter, M. Gyarmati, and U. Schilcher, "An inhomogeneous spatial node distribution and its stochastic properties," in *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, MSWiM '07, 2007.

[20] M. Gyarmati, U. Schilcher, G. Brandner, C. Bettstetter, Y. W. Chung, and Y. H. Kim, "Impact of random mobility on the inhomogeneity of spatial distributions," in *IEEE GLOBECOM 2008.*

[21] N. Aschenbruck, A. Munjal, and T. Camp, "Trace-based mobility modeling for multi-hop wireless networks," *Comput. Commun.*, vol. 34, May 2011.

[22] The Scenario Generator, "http://isis.poly.edu/ qiming/scengen/index.html."

[23] Scalable Network Technologies, "Qualnet 4.0."

[24] J. Yoon, M. Liu, and B. Noble, "Random waypoint considered harmful," in *INFOCOM 2003.*, 2003.

[25] C. Tuduce and T. Gross, "A mobility model based on wlan traces and its validation," in *IEEE INFOCOM 2005.*

[26] C. Bettstetter, "Smooth is better than sharp: A random mobility model for simulation of wireless networks," in *MSWIM '01*, 2001.

[27] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," *Mobile Computing, IEEE Transactions on*, vol. 2, no. 3, pp. 257 – 269, 2003.

[28] M. Piórkowski, N. Sarafijanovic-Djukic, and M. Grossglauser, "On clustering phenomenon in mobile partitioned networks," in *Proceeding of the 1st ACM SIGMOBILE workshop on Mobility models*, MobilityModels '08, pp. 1–8, 2008.

[29] L. Harfouche, S. Boumerdassi, and E. Renault, "Weighted social manhattan: Modeling and performance analysis of a mobility model," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, 2010.