# On the Symmetry of User Mobility in Wireless Networks

Bruno A. A. Nunes[1] 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 analyzed WLAN–, GPS–, and synthetic traces that record mobility in a variety of network environments. We observe that from a macroscopic level, human mobility is *symmetric*. In other words, the number of users that move from point $A$ to point $B$ approximates the number of users that go in the opposite direction, i.e., from $B$ to $A$. We show that this type of symmetry is more accentuated in synthetic mobility models, in particular, in random way-point mobility. We also study the direction of movement which also exhibits symmetric behavior in both real– as well as synthetic mobility. Additional contributions of our work include metrics to quantify mobility symmetry. We conclude the paper with a discussion of possible applications of our results in mobile networking.**

## I. Introduction

Node mobility is a key factor in the design and performance evaluation of mobile networks and their protocols and mobility characterization has attracted considerable attention from the networking research community. Evaluation studies of early mobile networks and their protocols used most of the time "synthetic" mobility models such as random walk, Brownian motion, and the random way-point (RWP) model [1], just to mention a few. The RWP, in particular, has been one of the most used mobility models for evaluating mobile networks which motivated several studies that scrutinized its behavior, identified a number of undesirable features [2], as well as proposed variations to improve its behavior.

More recently, motivated in part by the problems associated with the RWP and recognizing the importance of employing more realistic mobility scenarios when designing and evaluating mobile networks, there has been considerable interest in using real mobility traces and developing models that reflect real mobility. Crawdad [3], is an example of an initiative to make real mobility traces widely available to network researchers.

In this paper, also motivated by the trend towards employing real mobility to design and evaluate wireless networks, we study different types of traces obtained by recording user mobility. Our goal is to identify patterns, extract features, and define metrics to characterize the *spatial* behavior of human mobility. As a result, we identify an interesting characteristic exhibited by human mobility that, to the best of our knowledge, have not yet been revealed in previous studies. We work with a number of real traces that record user mobility in infrastructure-based networked environments (i.e., wireless LANs) as well as GPS positioning traces. We show that, from a macroscopic level, human mobility is *symmetric*. In other words, the number of users that move from point $A$ to point $B$ approximates the number of users that go in the opposite direction, i.e., from $B$ to $A$. We also show that this type of symmetry is more accentuated in synthetic mobility models, in particular, in random way-point mobility. Additionally, we study the direction of movement which also exhibit symmetric behavior in both real– as well as synthetic mobility. In order to quantify the degree of symmetry exhibited by a given mobility scenario, we define a new metric we call *coefficient of symmetry*. Finally, we hope that this work will generate input for the development of new and more realistic models, and that the new metrics presented here will aid in the evaluation of these models.

The remainder of this paper is organized as follows. Section II describes the mobility traces we studied, including wireless LAN, GPS, and synthetic traces. In Section III, we describe our methodology, define the metrics we used, and present our results. Section IV discusses the implications of mobility symmetry in areas such as design of forwarding algorithms, mobility management, mobility prediction, and the design of mobility models. We finally present some related work in Section V.

## II. Mobility Traces

In this section we describe the mobility traces we used in our study. Here, we make a distinction between different types of traces and define *macro*– and *micro*–*mobility*. In *macro-mobility* traces, the resolution of node positions is coarser grained; this is the case of WLAN traces or cell phone traces, where the exact position of a node is unknown and is instead represented by the position of the access point with which the node is associated (in WLANs) or the cell in which the node is presently located (in cellular networks). In *micro-mobility* traces, a node's location is represented with finer grain resolution; this is the case of GPS traces, where positioning information is given by latitude and

[1]Financial support was granted by the CAPES Foundation Ministry of Education of Brazil, Caixa Postal 250, Brasilia - DF 70040-020 Brazil.

978-1-4577-0351-5/11/$26.00 ©2011 IEEE

longitude, and typically has a resolution of meters.

We considered three types of mobility traces in our study, namely: infrastructure-based scenarios (WLANs); ad-hoc scenarios (using GPS positioning), and synthetic scenarios (using RWP mobility). Below we present the traces in these different categories in more detail.

### A. WLAN Traces

Table I summarizes the WLAN traces in terms of number of users, number of Access Points (APs), and duration of the trace. The first 3 WLAN traces were collected in university campus environments: Dartmouth and Stanford, available at [3] and the MIT trace [4].

| Trace | # users | # APs | Duration |
|---|---|---|---|
| Dartmouth | 9480 | 623 | 1 year (2003) |
| Stanford | 74 | 21 | 12 weeks |
| MIT | 1366 | 173 | 1 day |
| Rio | 120 | 17 | 1 week |

Table I
WLAN traces description.

Since our goal is to study human mobility in general, not only in the context of a campus environment, we collected a fourth trace recording user mobility in a public WLAN network deployed in an urban environment. The *Rio* trace was collected during the first week of April 2010 recording user mobility in the *Digital Orla Project* [5] network deployed along Ipanema Beach, Rio de Janeiro, Brazil. At the time the trace was collected, the network consisted of 17 Cisco APs mounted on light poles along the beach's boardwalk, providing network access in the boardwalk and nearby streets. The APs were configured to transmit a *syslog* message every time a client associated or disassociated. The resulting trace reflects one week of network usage; a total of 120 distinct clients who performed at least 1 transition between APs were recorded. Clients use a variety of devices to connect to the network including laptops and smart phones.

### B. GPS Traces

The focus here was to study user mobility with finer-grained resolution. WLAN traces represent "on-off" type behavior, where users connect to an AP, stay connected for a given period of time, disconnects, stays disconnected for a period of time, reconnects again to some other AP, and so on. GPS traces, represent mobility in a more continuous fashion where nodes roam around a given area, and connection between nodes and an AP or another node is a function of the distance between the nodes and the transmission range of their communication devices. Table II summarizes the GPS traces we used in our study in terms of number of users, trace duration, and the period between GPS samples.

| Trace | # users | Duration | Samples |
|---|---|---|---|
| Quinta | 98 | 900 s | 1 s |
| KAIST | 78 | 5000 s | 10 s |

Table II
Description of the GPS traces studied.

*Quinta*, refers to the "Quinta da Boa Vista Park" trace, presented in [6]. It is a GPS trace collected at a park in the city of Rio de Janeiro, Brazil, that has many trees, lakes, caves and trails, and holds the National Museum and the city Zoo. The KAIST trace, also available at [3], on the other hand, is a GPS trace collected on a university campus environment.

### C. Simulated Traces

Two synthetic traces were generated using the RWP mobility model. We tried to simulate the two scenarios described by the traces in Section II-B using RWP mobility. Our goal here is to compare characteristics of synthetic mobility against real mobility, especially when the former is inspired by the latter. In our simulations, we set the velocity range in a way that the average velocities would match the ones measured in the Quinta and KAIST traces. The same was done for average pause time and the dimensions of the area covered by the traces. Velocity and pause time were chosen according to a *uniform distribution* whose average is computed as described above [2]. Table III summarizes the parameters used to generate the synthetic traces.

| Parameter | Quinta_RWP | KAIST_RWP |
|---|---|---|
| Avg. Speed | 1.2 m/s | 0.72 m/s |
| Avg. Pause | 3.6 s | 17 s |
| Area | 840 m X 840 m | 5000 m X 5000 m |
| Duration | 900 s | 5000 s |
| # nodes | 98 | 78 |

Table III
Simulation parameters.

In order to study transitions we divided the area into square cells. We assumed 100 meters of range for a Wi-Fi device and set the cell to be the square inscribed in a 200 meters diameter circle. That gives approximately a square cell of size 140 X 140 meters.

## III. Dataset Analysis

In this section, we present the methodology we used to analyze the traces and the results of our study. We start by defining the metrics and parameters employed in our analysis.

### A. Definitions

**Transition** is the event caused when a node moves from one cell to another. When analyzing the GPS traces, transitions are detected when nodes move between neighboring cells. In WLAN traces, transitions happen between any cells.

**Transition Matrix** (TM) $A$ is a square matrix of $N$ x $N$ elements where $N$ is the total number of cells in the system. Every element $a_{ij}$, denotes the number of transitions observed from cell $i$ to cell $j$. Since we do not consider transitions when nodes move within the same cell, the main diagonal is zero.

**Time-Aggregated Transition Matrix** For a given window of time, we count how many transitions were observed for a given user. The larger this window, the

[2]For more details on how these parameters are derived from real traces refer to [6].

more we aggregate transitions in time. The window can be *sliding* or *fixed* over time. If we choose a sliding window, we define a window size $W$ and a starting time instant, say $t_{1,i}$, when the first transition for user $i$ occurred. We then count how many transitions occurred from $t_{1,i}$ to $t_{1,i} + W$. After that, the beginning of the window slides to the time of the next transition, say $t_{2,i}$ and the transitions between $t_{2,i}$ to $t_{2,i} + W$ are counted. This is done for all transitions for the duration of the trace. In the case of a fixed-window, the trace is divided in slots of size $W$ and we count the transitions that happen inside every slot. We used the fixed-window approach to compute aggregated transitions in the results presented below.

**User-Aggregated Transition Matrix** The TM can be computed for each user in the system separately, called here individual transition matrix (iTM), or for a number of users $n$, an *n-aggregated transition matrix* (n-aTM). For a time window of fixed size, user aggregation refers to how many users we consider when computing the TM.

**Symmetric Transition Matrix** $A$ is considered symmetric when $a_{ij} \approx a_{ji}, \forall i, j \in \{1, ..., N\}$.

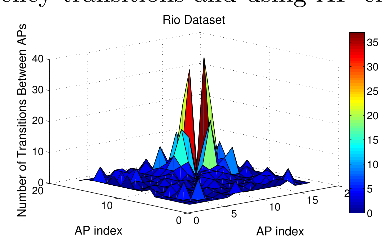
**Direction Angle** is the counter-clockwise angle formed between the X-axis (0°) and the line connecting two consecutive location samples in the trace.

### *B. Transition Symmetry*

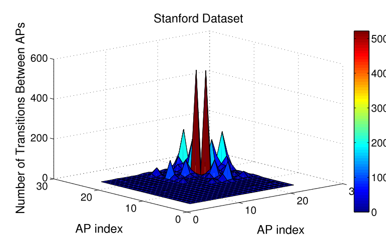
Figures 1(a), 1(b), and 1(c) show the aggregated TMs for the Rio, Stanford, and Dartmouth datasets, respectively. These TMs consider transitions within 24 hours and aggregate all users for the duration of the trace according to Table I. The MIT TM is not shown due to space constrains but shows similar behavior.

The aggregated TMs show a clear symmetry in relation to the main diagonal. One might argue that the symmetric behavior could be a consequence of the *ping-pong effect* reported in [7], commonly seen in WLAN traces. This phenomenon occurs when a mobile node is in range of two or more APs, and because of the variation in signal strength, this node might go back and forth associating and disassociating with the different APs in range. It is easy to identify this behavior if the session time is in the order of seconds, but when the gap between transitions is in the order of minutes, this task becomes non-trivial. In order to mitigate this behavior, we filtered high frequency transitions, only accounting for transitions that happened more than 1, 5, 10 and 15 minutes away from each other. As we increased the time gap, the number of transitions per node decreased as expected. However, it did not affect the symmetric behavior of the TM. For that reason, we set the *ping-pong filter* to 10 min for all the results presented here.

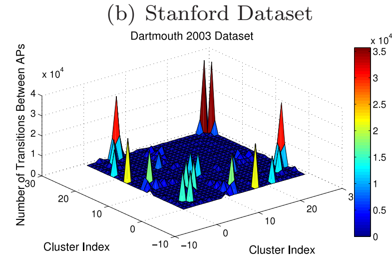
In order to further explore the reasons for the TMs' symmetric behavior, we considered clusters of APs such that if a transition occurred between APs in the same cluster, either due to node movement or the *ping-pong* effect, that would not count as a transition. In other words, only transitions between clusters were considered. We use AP clustering in the Dartmouth trace, which was the one with the largest number of APs. A map of the Dartmouth campus was used to decide how to cluster the APs: we used the known locations of the APs to overlay them on the map; we then clustered the APs according to proximity and area of coverage (instead of using clusters of similar size and number of APs). The Dartmouth TM represented in Figure 1(c) is a result of organizing the campus APs in 30 clusters. Evidently, symmetry persisted even after filtering out high frequency transitions and using AP clusters.




(a) Rio Dataset




(b) Stanford Dataset




(c) Dartmouth Dataset.

Figure 1. Transition matrices for the WLAN traces.

The explanation for the symmetric behavior is illustrated in Figure 2. If $user_1$ takes the path from $AP_2$ to $AP_1$ to $AP_3$ and $user_2$ goes from $AP_2$ to $AP_3$ to $AP_1$, then the resulting transition matrix is symmetric. The more we aggregate users and/or the longer we observe their movement, the higher is the chance of observing movement in opposite directions.

In order to understand and be able to quantify symmetry as we observe user mobility collectively and over time, we define a metric called *coefficient of symmetry*, $\alpha$, where $0 \leq \alpha \leq 1$, as defined in Equation 1:

$$\alpha = \frac{1}{T} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} |a_{ij} - a_{ji}| \tag{1}$$

where $T$ is the total number of transitions in matrix $A$. Thus, the closer $\alpha$ is to 0, the higher the symmetry. $\alpha = 1$ means that $A$ does not have any transitions in symmetric positions. This metric can also be seen as the percentage of transitions happening in symmetric positions. For example, if $\alpha = 0.2$, it means that 80% of the transitions are in symmetric positions. $\alpha$ can be computed for each node individually, or for the entire network from a TM aggregated over all users.

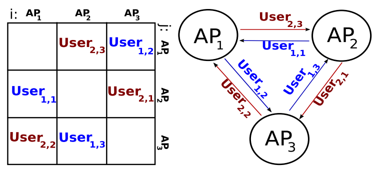
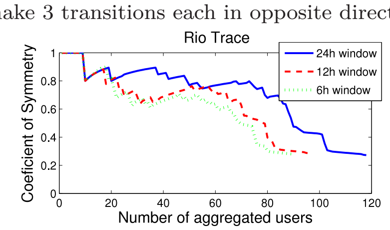
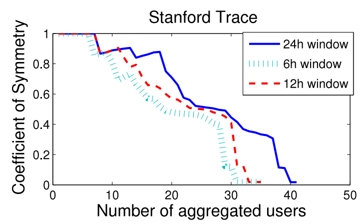


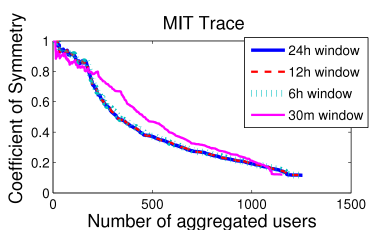

Figure 2. Transition matrix built for a network of 3 APs, after two users make 3 transitions each in opposite directions.



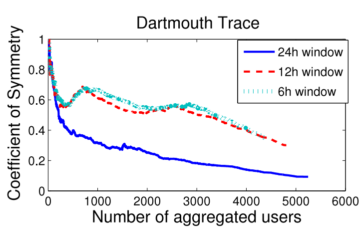

(a) Rio Dataset




(b) Stanford Dataset




(c) MIT Dataset




(d) Dartmouth Dataset

Figure 3. Metric of symmetry $\alpha$ for different number of aggregated users, under different time aggregation windows for (a) Rio, (b) Stanford, (c) MIT and (d) Dartmouth datasets.

Figure 3 shows $\alpha$ for all 4 WLAN traces studied, varying the number of aggregated users and aggregating transitions in time for 3 different window sizes: 24–, 12–, and 6 hours. The X-axis is how many users were aggregated to compute $\alpha$, and users are ordered in decreasing order, according to their individual $\alpha$ (computed using their iTM).

For the traces with larger number of users, when we decrease the aggregation time window, the TM becomes less symmetric (i.e., $\alpha$ increases), as we can observe from the Dartmouth TM (Figure 3(d)). This behavior is not as evident in the MIT trace (Figure 3(c)). This is due to the fact that all the APs are in the same building which greatly contributes to the ping-pong effect, making it much harder to filter. Therefore, decreasing the aggregation window does not affect the results as much. This happens because users' transitions, in this case, occur with great frequency, always falling inside the minimum window used (6 hours). To confirm this hypothesis we used an additional window of 30 minutes for the MIT dataset. Only then we were able to notice the difference, as the 30 minutes curve went over the other curves as we aggregated more users.

In the case of the traces with fewer users, time aggregation appears to have an opposite effect on the symmetry, as we can see in Figures 3(a) and 3(b). The smaller the time window is, the less transitions are accounted in the TM. In the case of those two traces, there is a preference for a few specific APs (more popular APs appear as the spikes in the TMs on Figures 1(a) and 1(b)). The transitions not accounted for due to the smaller time window are more likely to be taken from the most popular APs, what would decrease the difference amongst the cells in the matrix. This would increase the overall symmetric pattern, contributing to a lower $\alpha$ and making the $\alpha$ curve converge faster to lower values as we aggregate more users.

In Figure 4 we plot the coefficient of symmetry, $\alpha$, for the GPS and RWP traces. For the RWP traces, we plot the average aggregated $\alpha$ for 15 different random seeds. The first interesting observation is that $\alpha$ for the synthetic traces is similar to $\alpha$ for the corresponding real traces. The differences are due to the fact that in real environments, users tend to favor some areas (e.g., they walk on roads), limiting the degree of freedom of their movement. On the other hand, for the synthetic traces, users tend to disperse more uniformly increasing the chances of the number of transitions between a cell and its neighbors to be similar for every cell. For the KAIST trace, in particular, we notice that nodes tend to concentrate even more than in the Quinta trace, visiting less cells overall. That explains KAIST's higher symmetry, even with less users. This limited number of visited cells, in the KAIST trace, increases the relation user/cells increasing the chances of random transitions falling in symmetric positions.

### *C. Direction Angle Symmetry*

Direction of movement provides additional information on the spatial behavior of human mobility. Figure 5 shows the distribution of direction angles which represent relative frequency of directions users take when moving as reported in the GPS– and synthetic traces. The bin size used in these plots is 1°. We observe a symmetric–, close to uniform distribution of direction angles for both of the synthetic traces (Figures 5(b) and 5(d)), where it is not possible to identify any significant preference for a given direction.

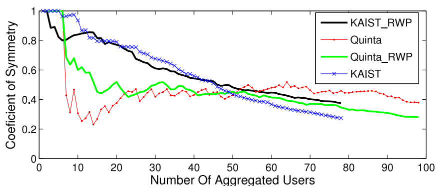
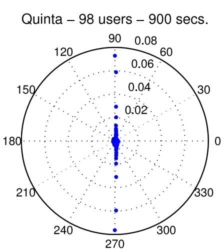
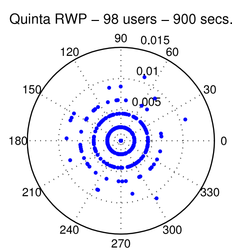



Figure 4. $\alpha$ per number of aggregated users, under a 24h aggregation window.

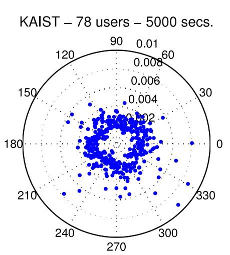
For the KAIST trace, in Figure 5(c), it is possible to notice the concentration of points over a symmetric shape, like a circle. Despite some few outliers, it is also not possible to identify any significant preferred angle. That is not true for the Quinta trace in Figure 5(a). This trace has a peculiar symmetric shape with respect to 180 degrees, for the direction nodes take. This can be explained by considering the physical terrain and the fact that users mostly move only along roads and tracks which limit the possible movement directions.



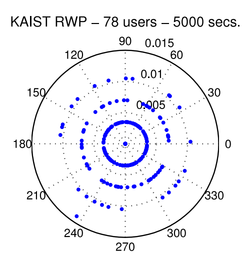

(a) Quinta Dataset




(b) Quinta RWP Dataset




(c) KAIST Dataset




(d) KAIST RWP Dataset

Figure 5. Distribution of direction angles for (a) Quinta, (b) Quinta RWP, (c) KAIST and (d) KAIST RWP datasets.

The KAIST trace, on the other hand, was collected on a much larger area which corresponds to the KAIST campus, with several roads and pathways. Because it was derived from the real KAIST trace, the KAIST_RWP trace also presents this characteristic. This is what is shown in Figure 6 which plots the Empirical Cumulative Distribution Function (ECDF) of the direction angles for the GPS– and synthetic traces. We observe that the KAIST_RWP trace is able to reflect closely the statistical properties of the trajectory angle. The differences between Quinta and Quinta_RWP traces are mainly due to movement restrictions (obstacles and paths).

## IV. Discussion

Despite the large body of work that focus on studying real traces, to the best of our knowledge, our work is the first to observe and quantify spatial symmetry in user mobility in terms of cell transitions and angle of movement. The only other work that briefly alludes to direction angle symmetry is [8].

Using maps, social interaction,"hot spots", and other factors that limit and/or influence user mobility (e.g., what is the next destination) are effective strategies to "bias" synthetic mobility models in order to obtain behavior that is closer to reality. Along those lines, the proposed TM can be used to identify the "popularity" of the different regions of the network (including, e.g., hot spots) and based on this information specify the probability distributions used by nodes to select their next destination in synthetic mobility models.

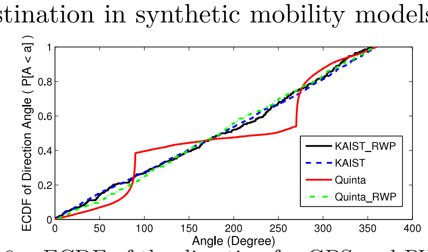



Figure 6. ECDF of the direction for GPS and RWP traces.

Another interesting application of our results is the use of the direction angle symmetry as a way to compare mobility traces. The distributions of direction angles in Figures 5(a) and (b) show very different behavior when comparing the synthetic trace against the real one. By using strategies to approximate synthetic mobility models to real mobility, direction angle symmetry can be used to quantify how close the resulting models get to real mobility, and thus how effective these "approximation" strategies are. We are currently going a step further and trying to draw a transitive relationship between direction angle symmetry, spatial mobility behavior, and performance of core networking functions such as routing and forwarding.

Mobility or location prediction is yet another area in which our results can find interesting applications. For example, in [9], a Markovian predictor is used to predict user location based on a given mobility trace (in the paper, the Dartmouth trace was used). It was observed that for nodes that did not have enough historical information recorded in the trace, the performance of the predictor was not adequate. To mitigate this problem, we could use the coefficient of symmetry, $\alpha$, of a given node to artificially "re-create" the node's past historical information.

## V. Related Work

Several empirical studies have focused on characterizing mobility behavior in WLANs. For example, the work reported in [10] studies a campus WLAN trace and extracts statistics such as the number of associations per time per AP as well as session duration. Their goal was to generate input for capacity planning. The MIT WLAN trace was used in [4], where statistics on traffic and AP associations were presented; the work also introduced important metrics such as *prevalence* and *persistence*. The study reported in [11] explored the Stanford traces but focused on traffic– and application-layer issues rather than mobility. Nonetheless, empirical distributions were derived for the number of users per AP and the maximum number of hand-offs per AP. Campus WLAN usage was further studied in [12] using the Dartmouth traces. Some of their findings include the fact that users tend to persist at a single location for longer and that different applications had different mobility characteristics. Dartmouth's, MIT's, and two other traces were studied further in [13] which presented ECDFs for time spent in each AP, number of associations per user, and percentage of APs visited. Moreover, metrics such as clustering coefficient, degree of separation between nodes, and disconnection ratio were also reported.

Trace-driven approaches aimed at building realistic mobility models have been the focus of many research efforts. In [14], a weighted waypoint model is proposed based on a university campus mobile scenario containing 5 possible locations. The transition probabilities were set based on a survey conducted on site, where people were asked about their mobility patterns on campus. A transition matrix was built as a result of this survey; however, it did not exhibit any symmetry. We believe this was the case because the transitions did not take into consideration intermediate locations. In [7], a framework for building trace-driven mobility models was proposed. It uses a map-based model where transition probabilities are computed from route weights and depend on origin, destination, current and previous locations, all extracted from the Dartmouth trace. In a more recent study [15], a model where the next waypoint depends on cell popularity and distance from current position is proposed.

## VI. Conclusions and Future Work

This paper explores the spatial behavior of human mobility through a variety of mobility traces collected in different network environments. As a result, we identify characteristics exhibited by human mobility that, to the best of our knowledge, have not yet been revealed in previous studies. We used a number of real traces that record user mobility in wireless LANs as well as GPS traces. Additionally, we also investigated synthetic mobility, in particular, through traces generated using the RWP model. We show that, from a macroscopic level, human mobility is *symmetric*. In other words, the number of users that move from point $A$ to point $B$ approximates the number of users that go in the opposite direction. We also show that this type of symmetry is more evident in synthetic models such as RWP. Additionally, we study the direction of movement which also exhibit symmetric behavior in both real– as well as synthetic mobility. In order to quantify the degree of symmetry exhibited by a given mobility scenario, we define a new metric we call *coefficient of symmetry*. We conclude by discussing the implications of our results in different areas of mobile networking.


## Acknowledgements

Authors would like to thank Bruno Sanso, Athanasios Kottas and Kumar Viswanath for their invaluable contribution to this work. This work was partially supported by NSF grant CCF-091694 and a US Army-ARO MURI grant.


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

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

[4] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," in *Proceedings of MobiSys 2003*.

[5] Orla Digital Project, "http://www.ravel.ufrj.br/orla/ipanema/cobertura3.php," COPPE/UFRJ 2010.

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

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

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

[9] L. Song, D. Kotz, R. Jain, and X. He, "Evaluating location predictors with extensive wi-fi mobility data," in *2004 IEEE INFOCOM*.

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

[11] D. Tang and M. Baker, "Analysis of a local-area wireless network," in *Proceedings of the MobiCom 2000*.

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

[13] W. jen Hsu and A. Helmy, "Impact: Investigation of mobile-user patterns across university campuses using wlan trace analysis," 2005.

[14] W.-j. Hsu, K. Merchant, H.-w. Shu, C.-h. Hsu, and A. Helmy, "Weighted waypoint mobility model and its impact on ad hoc networks," *SIGMOBILE Mob. Comput. Commun. Rev.*

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