This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2009 proceedings.

SLAW: A Mobility Model for Human Walks

Kyunghan Lee (KAIST), Seongik Hong (NCSU), Seong Joon Kim (NCSU),
Injong Rhee (NCSU) and Song Chong (KAIST)

***Abstract*—Simulating human mobility is important in mobile networks because many mobile devices are either attached to or controlled by humans and it is very hard to deploy real mobile networks whose size is controllably scalable for performance evaluation. Lately various measurement studies of human walk traces have discovered several significant statistical patterns of human mobility. Namely these include truncated power-law distributions of flights, pause-times and inter-contact times, fractal way-points, and heterogeneously defined areas of individual mobility. Unfortunately, none of existing mobility models effectively captures all of these features. This paper presents a new mobility model called *SLAW (Self-similar Least Action Walk)* that can produce synthetic walk traces containing all these features. This is by far the first such model. Our performance study using using SLAW generated traces indicates that SLAW is effective in representing social contexts present among people sharing common interests or those in a single community such as university campus, companies and theme parks. The social contexts are typically common gathering places where most people visit during their daily lives such as student unions, dormitory, street malls and restaurants. SLAW expresses the mobility patterns involving these contexts by fractal waypoints and heavy-tail flights on top of the waypoints. We verify through simulation that SLAW brings out the unique performance features of various mobile network routing protocols.**

I. Introduction

In mobile networks, the performance of networking applications highly depends on the movement patterns of wireless device holders. As wireless devices are mostly attached to humans, understanding human mobility patterns leads to an accurate performance prediction of protocols used for these networks. These patterns can be used to simulate mobile networks whose mobility governed by humans. Because of the difficulty in deploying real mobile networks big enough for the meaningful performance evaluation, simulation has been primarily used for performance evaluation. Therefore, mobility models reproducing essential mobility patterns of humans are important components of those simulation tools. This paper addresses this issue.

There have been several recent studies [9], [17], [35] reporting the discovery of fundamental statistical features of human mobility from real traces of human mobility such as GPS traces of human walks in various locations [35] and cellphone location tracking [17], recordings of wireless device associations with their access points [27] and tracking of bank notes [9]. They collectively report the statistical behaviors of human mobility over various scales of time and space. The followings summarize those features.

- **(F1) Truncated power-law flights and pause-times.** It is shown that the lengths of human *flights* which are defined to be straight line trips without directional change or pause (or lines between two consecutive waypoints) have a truncated power-law distribution [9], [17], [35]. Several studies (e.g., [35] [27]) also show that the pause time distributions of human walks follow a truncated power-law distribution.
- **(F2) Heterogeneously bounded mobility areas.** Gonzalez et al. [17] report that people mostly move only within their own confined areas of mobility and different people may have widely different mobility areas.
- **(F3) Truncated power-law inter-contact times (ICTs).** The distribution of *inter-contact times*, that is, the times elapsed between two successive contacts of the same persons can also be modeled by a truncated power law distribution [11] which consists of a power-law head followed by an exponentially decaying tail after a certain characteristic time [25].
- **(F4) Fractal waypoints.** From the analysis of the GPS traces of human walks, we report in [34] that the waypoints of humans can be modeled by fractal points. This implies that people are always more attracted to more popular places.

All these patterns ($F1-F4$) are intrinsically related to each other. We show in this paper that fractal waypoints ($F4$) induce power-law flights ($F1$). It is also shown in [20] that heavy-tail flights within a confined area ($F2$) result in truncated power-law ICTs ($F3$). It is also important that these patterns are strong performance determinants of mobile networks such as Delay-Tolerant Networks (DTN). For instance, ICTs are a very important factor to DTN routing as ICTs decide the delays before meeting a distant destination; so short ICTs imply short routing delays. Fractal waypoints express some social contexts such as common gathering places among people that share common interests or those in the same community. These contexts are important as they influence contact probability among people. Therefore, for an accurate performance evaluation of mobile networks, synthetic mobility traces to represent these patterns realistically.

Unfortunately, none of existing mobility models for mobile networks produce synthetic human mobility traces that possess all these statistical features. For instance, no mobility models, except Truncated Levy Walk [35], explicitly model heavy-tail flights. Random mobility models such as Truncated Levy Walk, Random Way Point and Brownian Motion do not produce fractal waypoints nor do they model heterogeneous mobility ranges of individuals. Although many models (e.g., [7], [16], [21], [23], [27], [29]) explicitly model hot spots or grouping effects, they do not produce heavy-tail flights or

978-1-4244-3513-5/09/$25.00 ©2009 IEEE

power-law ICTs. Those producing heavy-tail ICTs (e.g., [8], [25]) also do not necessarily model heavy-tail flights. Table I documents the statistical features modeled by many existing mobility models.

In this paper, we present a new mobility model, called *Self-similar Least-Action Walk* (SLAW), that produces synthetic mobility traces containing all those features. This is the first such model. In developing SLAW, we heavily rely on our GPS traces of human walks [34] including 226 daily traces collected from 101 volunteers in five different outdoor sites. In particular, many of these traces are gathered among people sharing the common interests such as students in the same university campuses and tourists in a theme park. By faithfully representing the patterns present in these traces, SLAW represents inherent social contexts among walkers manifested as common gathering places and walk patterns therein. By also modeling both power-law flights and fractal waypoints, SLAW can also express the regular as well as spontaneous trip patterns present in the daily mobility of humans. People typically keep a routine of visiting the same places every day such as going to an office, but at the same time, make irregular trips. It is not the case where people would always randomly choose places to visit and visit them in a random order. While other work [14], [22], [26] exists in expressing the regularity of daily trip patterns of humans, none of the existing work reflects realistic statistical patterns appearing in real human walk traces. Expressing this regularity and spontaneity realistically by modeling a human movement decision mechanism is one of the key contributions of our work.

To measure how these mobility patterns influence the performance of mobile network protocols, we study the performance of DTN routing under SLAW and various other mobility models. Our study indicates that SLAW realistically brings out the unique performance features of various DTN routing protocols. Especially, compared to random mobility models, it provides a clear performance differentiation between memoryless and memory-full protocols where memory-full protocols utilize past contact history information among nodes to predict the future contact probability. Examples of these protocols are abundant (e.g., [3], [30], [37]).

SLAW can be an important tool for emulating human walk behaviors. The applications of our work does not stop at mobile network; it can be applied to accurate urban planning, traffic forecasting and biological and mobile virus spread analysis where emulating human mobility is essential. SLAW is also as simple to work with as random mobility models such as RWP and BM because it needs only a small number of input parameters such as the size of the walk-about area, the number of walkers, and the Hurst value used for generating fractal waypoints. Note that many sophisticated models (e.g., [7], [16], [27], [29]) require detailed information such as mobility transition probability tables and location and sizes of hot spots which are only available from real traces. SLAW does not require any real walk traces for generating synthetic traces.

The remainder of this paper is organized as follows. Section II has related work . Section III describes SLAW in detail.

Category	Features / Models	$F1$	$F2$	$F3$	$F4$
Random	RWP, RD	N	N	N	N
	BM(RW)	N	N	Y	N
	TLW [35]	Y	N	Y	N
Random Variants	MWP [23]	N	N	N	N
	GM [28]	N	N	?	N
	RPGM [21]	N	N	?	N
Geographic	Freeway [2]	N	N	?	N
	Manhattan [2]	N	N	?	N
	OM [24]	N	N	?	N
Social	Dartmouth [27]	N	N	?	N
	CMM [29]	N	N	?	N
	ORBIT [16]	N	Y	?	N

TABLE I
EXISTING ROUTING MODELS CAN BE CATEGORIZED INTO FOUR GROUPS. TYPICAL MODELS IN EVERY CATEGORY HAVE BEEN LISTED. NONE OF THE EXISTING MODELS HAVE ALL THE CHARACTERISTICS OF HUMAN WALKS. '?' MEANS THAT IT IS UNCLEAR FROM THE MODEL DESCRIPTION.

Section IV discusses the validation of SLAW and our routing protocol study. Section V has the conclusion.

II. RELATED WORK

Table I shows four categories of existing models and each model. Random Way Point (RWP), Random Direction (RD), Brownian Motion (BM) or Random Walk (RW) and Truncated Levy Walk (TLW) [35] are random mobility models. In random mobility, each waypoint is chosen randomly based on some probability distribution. Markovian Way Point (MWP) [23] and Gauss-Markov (GM) model [28] are slight variants of random models as they implement some Markovian transition probabilities among waypoints or prohibit unrealistic abrupt velocity changes. In RPGM [21], mobile nodes form several groups each of which contains one leader. All the members of a group move along their leader.

Other models incorporate geographical constraints or social contexts and collective behaviors. The Obstacle Model (OM) with geographical constraint [24] incorporates obstacles to emulate more realistic pathways of humans around obstacles using Voronoi diagrams. To emulate pathways the Freeway and Manhattan mobility models [2] also have been proposed, and they restrict the movements of mobile nodes to follow the pathways. Modeling hot spots is another way to represent collective human behaviors.

The Dartmouth model [27] estimates the locations and movement paths of mobile nodes from real data sets. Based on the estimated information on the users, hotspot regions and the transition probability for moving between hotspots are extracted. This model comes from real data but it requires a considerable amount of effort to generate the mobility model because hotspot locations and transition probability between hot spots must be given as input (instead of being generated). Thus, it is very hard to change the walkabout areas, the number of nodes and the locations of hot spots without any corresponding real data sets.

Some hot spot models [7] [29] use scale-free networks [6]. Using the preferential attachment theory [5], a set of *attractors* is established where attractors are either landmarks or nodes. For instance, Clustered Mobility Model (CMM) [29] is a good example. It first divides the simulation area into a number of

subareas and use them as attractors. Mobile nodes are assigned to a subarea using preferential attachment. The attractiveness of one area is determined by the current number of nodes assigned to that area.

ORBIT [16] randomly creates a specified number of clusters within a given area and each node are assigned to a subset of clusters. A node moves only among its assigned clusters. For movements between and within clusters, it is completely random. Like SLAW, ORBIT explicitly models the heterogeneously bounded walkabout areas of different people (i.e., $F4$). However, ORBIT does not capture fractal points and the choices of next destinations are completely random. Thus, it lacks the regular patterns present in daily human walks.

III. SLAW: Self-similar Least Action Walk

A. SLAW Overview

Capturing both power-law flights and fractal waypoints in walk traces is non-trivial. Due to mutual dependency among various parameters such as the degree of self-similarity (Hurst parameter) and the characteristics of flight distributions, e.g., the power-law slope of the distribution, controlling them concurrently is also hard. SLAW adopts the following approach to the problem. It first generates fractal waypoints ($F4$) using a technique similar to a fractional Gaussian noise or Brownian Motion generation technique (fGn or fBm) (e.g., [32], [33]) over a 2-D image. SLAW leverages fundamental properties of fractal points to generate power-law flights ($F1$) on top of them. In this paper, we prove analytically that fractal points induce power-law gaps where *gaps* are defined to be inter-spacing among neighboring fractal points. Our analysis recovers algebraic relations between the Hurst parameter of fractal points and the power-law slope of the corresponding gap distributions.

It has been empirically known that the gaps among fractal points follow a power-law distribution [15], [31]. Typically gaps over multiple dimensions are measured by Delaunay triangulation. We experimentally verify that the line lengths of Delaunay triangles drawn on top of the waypoints extracted from real human walk traces have a power-law distribution and furthermore, their power-law slopes are almost identical to those of flight distributions from real traces. This strongly suggests that people plan their trips over known destinations (if we view waypoints as destinations) in a *gap-preserving* manner where they visit the nearby destinations first before visiting farther destinations. This trip planning is in fact a heuristic to the well-known traveling salesman problem whose objective is to minimize the total distance of travel and aligns well with the *least-action principle* of Maupertuis [13] where all objects are moving to the direction of minimizing their discomfort. In a human walk case, the discomfort is distance. It is intuitive that when visiting a place, people strive to reduce the distance of travel by visiting all the nearby destinations before visiting farther destinations unless some high priority events such as appointments force them to make a long distance trip even in the presence of unvisited nearby destinations. The least-action principle has led us to developing a trip planning algorithm called *Least Action Trip Planning* (LATP). We show that LATP can recover the same flight distribution observed in real traces within around 10% error margin.

LATP and fractal waypoints may satisfy $F1$ and $F4$. To complete the trace generation process while satisfying $F2$ and $F3$, we combine LATP with an individual walker model. To enforce heterogeneously bounded walkabout areas among walkers, SLAW develops an individual walker model to restrict the mobility of each walker to a predefined sub-section of the total area. It is done by selecting a subset of fractal waypoint clusters and restrict the movement of each walker to its own designated set of clusters. To add randomness, it also allows walkers to move out of their predefined walkabout areas occasionally with some controlled probability. We verify that this walker model combined with LATP and fractal waypoints generate power-law ICTs (satisfying $F3$).

More important, SLAW can realistically represent the regular patterns of human daily mobility. Since it is typical that people maintain a fixed routine of daily mobility such as going to an office, expressing these patterns are important. We verify later that the combination of fractal waypoints, LATP and the individual walk model effectively create this regularity as well as randomness and spontaneity in human mobility.

Below, we present our analysis that fractal points induce power-law gaps, describe LATP, and then present the new individual walker model. The description on our fractal point generation is omitted due to the space limit and the detailed description can be found in [34]. The technique is a simplified version of fBm [32].

B. Fractal Points and Power-law Gaps

We first examine the relation between one-dimensional fractal points and power-law gap and later we generalize our analysis over to two dimensional cases through simulation. Consider a process of dispersing a set of points over one dimensional space. Let $Y(x)$ be the number of points over a line interval (x and $x+\Delta$). In other words, $Y(x)$ is a point count sequence over a small interval Δ. The self-similarity of point counts $Y(x)$ can be manifested in several equivalent ways. First, the aggregated variance $v(m)$ of $Y(x)$, which is a variance of a new series by averaging the original series $Y(x)$ in nonoverlapping blocks of m, replacing each block by its mean, has an asymptotic form of $v(m) \sim m^{-\beta}$, $0 < \beta < 1$ as $m \to \infty$. The Hurst parameter can be expressed as $H = 1 - \beta/2$. Second, the power spectrum $S(f)$ of $Y(x)$ has $1/f$ noise around the origin, that is, $S(f) \sim f^{-\theta}$, $0 < \theta < 1$ as $f \to 0$.

The gap (interval) between two points can be measured as follows. We first make Δ small enough to hold at most one point and define the distance between any two immediately neighboring points as *gap*. Let $p(x)$ be the probability density function of a random variable x representing a gap among the fractal point process $Y(x)$.

Theorem 3.1: Fractal points over one dimensional space induce power-law gaps, that is, if $v(m) \sim m^{-\beta}$, $0 < \beta < 1$

as $m \to \infty$, then $p(x) \sim x^{-\alpha-1}$, $0 < \alpha < 1$ as $x \to \infty$. Furthermore, $\alpha + \beta = 1$.

Proof: Due to the space limit, we provide a proof sketch below. The proof is adopted from [4], [18].

From [12], we can have the following relation between $v(m)$ and the autocorrelation function $\rho(x)$ of $Y(x)$.

$$v(m) = v \cdot \Big(\frac{1}{m} + \frac{2}{m^2}\sum_{n=1}^{m}(m-n)\cdot\rho(n)\Big) \quad (1)$$

where v is the variance of $Y(x)$.

From Eq. 1, $\rho(x) \sim x^{-\beta}$ since $v(m) \sim m^{-\beta}$.

A correlation function $c(x)$ of $Y(x)$ is asymptotically the same as its auto-correlation function $\rho(x)$ and can also be represented as a conditional probability to have a point at x, given that a point occurs at a $x = 0$. $c(x = n\Delta)$ can be represented as follows [36].

$$c(\Delta) = p(\Delta) \quad (2)$$
$$c(2\Delta) = c(\Delta)p(\Delta) + p(2\Delta) \quad (3)$$
$$\cdots \quad (4)$$
$$c(n\Delta) = \sum_{k=0}^{n-1} c[(n-k)\Delta]p(k\Delta) + \delta_{n,0} \quad (5)$$

The Fourier transform $\mathcal{F}$ of the last equation gives a power spectrum of $Y(x)$. Thus, $S(f) = \mathcal{F}(c(x)) = \frac{1}{1-\mathcal{F}(p(x))}$. Let $\hat{P}(f)$ be an asymptotic function of $\mathcal{F}(p(x))$.

$$\hat{P}(f) = 1 - \frac{1}{S(f)} \quad (6)$$

Since $c(x) \sim x^{-\beta}$, $S(f) = \mathcal{F}(c(x)) \sim f^{\beta-1}$ as $f \to 0$. Since $0 < \beta < 1$, the power spectrum of $Y(x)$ has $1/f$ noise.

If we set $\alpha = -\beta+1$, then from Eq. 6, $\hat{P}(f) \sim 1-f^{\alpha}$. Note that the inverse Fourier transform of $1-Bf^{\theta}$, $0 < \theta < 1$ where B is a constant, is asymptotically $x^{-\theta-1}$. Therefore, $p(x) \sim x^{-\alpha-1}$ as $x \to \infty$. Since $0 < \alpha < 1$, $p(x)$ is asymptotically power-law and it proves $\alpha + \beta = 1$. ■

Although we can easily extend Theorem 3.1 over to multiple dimensions analytically, the definition of gaps over a multi-dimensional space is not intuitively clear. Gaite [15] uses Delaunay triangulation to measure the gaps over multi-dimensional spaces. We apply Delaunay triangulation on the waypoints we extracted from real traces [35]. The gap distributions over waypoints extracted from real traces can be seen in Fig. 3. Their power-law fitting is omitted and given in [34].

We measure the power-law slope (β) of the aggregated variances (β) of waypoints extracted from the GPS traces and the power-law slopes (α) of the CCDF of their corresponding gap distributions. We find that the values of $\alpha + \beta$ from real traces are close to 1.2. The margin of errors may arise from truncations caused by confined measurement areas as the truncations may significantly distort the power-law slope visible at the body of distribution. We also perform Delaunay triangulation on top of synthetically generated fractal waypoints using a simplified fBm technique [34] over a 2-D area and measure α and β. For each Hurst parameter value (H), we generate 10 synthetic waypoint maps. Figure 1 shows that $\alpha + \beta \approx 1$ with a similar margin of errors.

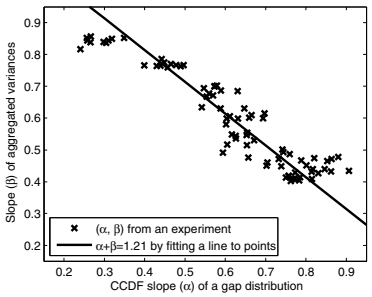


Fig. 1. α and β measured over synthetic 2-D waypoint maps.

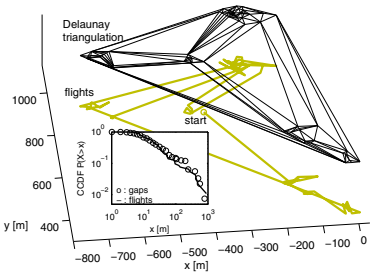


Fig. 2. Delaunay triangulation of waypoints extracted from one daily trace of KAIST. In the inset, the CCDFs of the line lengths from the triangles and the flights from the same trace are plotted. The CCDF of flights and gaps are closely matching.

C. Least-Action Trip Planning

In Section III-B, we showed that fractal points induce a power-law gap distribution. To find out how the power-law gaps are related to the actual human flights, we compare the flight distributions obtained from real traces and the gap distributions induced by the waypoints from the same traces. We find a strong similarity between these two distributions, especially in terms of their slopes and shapes. Figure 2 show that the Delaunay triangles on top of a daily trace of one participant, its line length CCDF forming the triangles and the CCDF of flights extracted from the same trace. We also perform the Delaunay triangulation on each individual trace and aggregate all the resulting triangle lines from all the traces. Figure 3 plots the resulting CCDFs. Their similarity with the corresponding flight distributions is strikingly impressive.

The similarity in the power-law slopes of gaps and flight distributions suggests interesting aspects about the order which a walker visits waypoints for a given set of waypoints. Obviously people are not conscious about gaps when they travel. However, as we can see from Figure 2, Delaunay triangles are formed among "neighboring" waypoints as Delaunay triangulation produces a planar graph where edges intersect only at their endpoints. From this, we conjecture that people

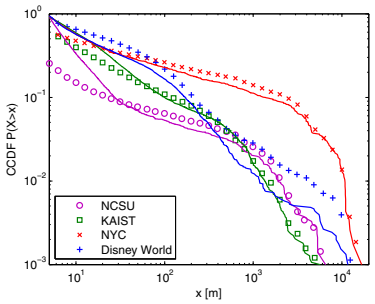


Fig. 3. Delaunay triangulation is performed on all individual daily traces. The line segment lengths in Delaunay triangles are aggregated and their CCDFs are plotted for different walkabout sites. The CCDF of flights obtained from the corresponding traces are also plotted for comparison.

might be minimizing the traveling distance. This makes sense intuitively when we are given a priori multiple destinations at different distances, we often strive to minimize the total distance of travel by first visiting nearby locations before visiting farther locations. In fact, this "greedy" way of trip planning is similar to a heuristic to the traveling salesman problem whose objective is to minimize the total distance of travel and aligns well with the least action principle of Maupertuis [13].

Algorithm 1 *Least action trip planning* (LATP) algorithm with a distance weight function $d^{-\alpha}$

V: set of all vertices (waypoints)
V': set of all visited vertices
$s \in V$: starting vertex
$c \in V$: current vertex
$c \Leftarrow s$
$V' \Leftarrow \{c\}$
while $V' \neq V$ **do**
  Calculate distances to all unvisited vertices, $d(c,v) = \|c - v\|_2$ for all $v \in V - V'$
  Calculate probability to move to all unvisited vertices,
$$P(c,v) = \frac{\left\{\frac{1}{d(c,v)}\right\}^{\alpha}}{\sum_v \left\{\frac{1}{d(c,v)}\right\}^{\alpha}} \text{ for all } v \in V - V'$$
  Choose a next vertex $v' \in V - V'$ according to the probabilities $P(c,v)$
  $c \Leftarrow v'$
  $V' \Leftarrow V' \cup \{c\}$
end while

To measure how much real human traces follow the least action principle, we measure the degree that people chooses their next destination based on the distance. This can be estimated as follows. We first measure the *flight-to-nearest-waypoint* ratio. For a given flight from x to y, suppose k is the nearest unvisited waypoint from x. The flight-to-nearest-waypoint ratio is the ratio of $||x-y||$ over $||x-k||$. We then define the *least-action criterion*: for a give flight, it tests if its flight-to-nearest-waypoint ratio is less than some threshold. Figure 4 plots the percentage of flights meeting the least-action criterion in real traces for all participants when the threshold ratio (r) is less than 2. On average, 53% of flights meet the criterion. However, if we consider that humans are less sensitive to the distance when next destinations are all nearby. So if we exclude those flights whose length is less than a short distance (say 30 meters), we get more than 87%

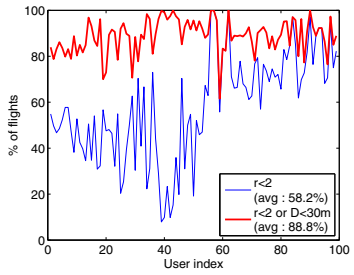


Fig. 4. For the GPS traces of 100 participants, we measure the percentage of flights meeting the least-action criterion. We also plot the case when we exclude those flights whose length is less than 30 meters.

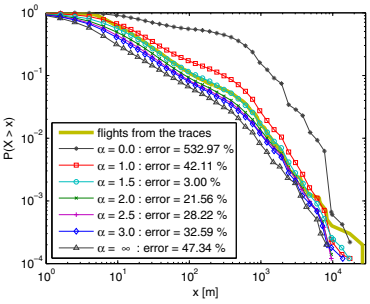


Fig. 5. Flight distributions obtained from LATP for different a values performed on top of waypoints extracted from the KAIST traces and the percentage difference between the sum of flights generated from LATP and that of flights from real traces

of flights meeting the criterion on average. We also tried other distances such as 20 meters and 10 meters and also varied r to a smaller ratio (1.5). All the measurements produce around 80% flights meeting the criterion. This indicates that most people in our traces use distance as an important metric for deciding the next waypoint, substantiating our conjecture.

We construct a new trip planning algorithm called *Least Action Trip Planning (LATP)* that given a set of waypoints to visit, decides the order in which a person visits. Algorithm 1 gives a pseudo-code of LATP. The algorithm selects a next waypoint to visit based on a weight function of $1/d^a$ where d is the distance from the current waypoint to an unvisited waypoint and a is a constant deciding the distance weight. If a is infinite, then the algorithm always chooses the nearest unvisited waypoint and if it is zero, then it randomly chooses the next waypoint. Figure 5 shows the flight distributions obtained from LATP for different a values performed on top of waypoints extracted from the KAIST traces and the percentage difference between the sum of flights generated from LATP and that of real flights from traces. In all cases, their difference is within a margin of around 10% when a is between 1 and 3. The figure shows that LATP generates traces very similar to the original traces.

D. Individual Walker Model

For a given input area S, our fractal waypoint generation generates a set W of the waypoints. We model an individual walker model that selects a subset of W and specifies the order

in which those selected waypoints are visited. When selecting these waypoints, we need to be careful. Fractal waypoints have a tendency of creating bursty clusters of various sizes dispersed over S. If waypoints are uniformly selected from W, then it is mostly like that all walkers will traverse through most clusters and will not satisfy $F2$. To satisfy $F2$, we need to assign different walkabout areas to different walkers and restrict each walker to move only around their designated areas. We develop a heuristic individual walker model to model these behaviors.

The algorithm works as follows. We first build clusters of waypoints by transitively connecting waypoints within a radius of 100 meters (typical WiFi outdoor transmission range). Let $C = \{c_i, i = 1, n\}$ be the cluster set and $|c_i|$ be the number of waypoints and T be the the total number of waypoints in S. We assign a weight $|c_i|/T$ to each cluster i. Then each walker k chooses 3 to 5 clusters (the exact number is controlled by the input) randomly from C with probability proportional to these weights – the cluster with a higher weight gets the higher chance to be picked. Let C_k be the set of the selected clusters. For each cluster, it uniformly chooses 5 to 10% of waypoints in each cluster randomly (the exact proportion is an input value); two different walkers are allowed to have the same waypoints. Let W_k be the set of waypoints walker k has selected from S. It also picks a starting waypoint from W_k from which it always starts its daily trip.

To add some randomness in his travel, each day, a walker k picks some additional waypoints W'_k as follows; it first chooses one new cluster c' randomly (ignoring weights) not in C_k and W'_k is the waypoints randomly picked from c' (about 5 to 10% of waypoints in c'). At the beginning of each day, walker k starts from its starting point and throughout the day, makes a one-round trip visiting all waypoints in $W_k \cup W'_k$ using LATP. (W_k does not change each day, but W'_k does). It uses a truncated power-law pause-time distribution to decide the amount time to stay at each waypoint. In the end of the day, it comes back to its starting point. The average pause time is adjusted so that the whole trip will end within a period of 12 hours.

Since each walker k always makes daily trips over a fixed set of W_k, its mobility area is bounded and also since each walker picks its set randomly, they tend to have different mobility area. In addition, we allow walkers to deviate from these waypoints by making them to pick some new waypoints additionally from the other clusters not in C_k. This allows walkers without any overlapping clusters occasionally meet, thus having some long ICTs. Those with overlapping clusters may have regular periodic contacts, depending on the transmission ranges or the time they arrive to the clusters.

We apply the cluster weights when selecting C_k to build some sense of community among all walkers. Because of fractal waypoints, some clusters are very large – so many walkers are likely to visit them. These clusters are emulating the common popular gathering places for all participants such as student union, dormitory, shopping, street malls or classrooms. Next section, we verify that SLAW with this individual walker model produces power-law ICTs ($F3$).

IV. Performance Evaluation

A. Simulation Setup

For validation, we run mobility simulations using various mobility models. We fix the simulation areas to be approximately the same as the measurement sites in [1]. The transmission range of each node is varied from 25 meters to 150 meters. If not explicitly said, it is set to 50 meters. 50 nodes are simulated for 200 hours and the first 50 hours of simulation results are discarded to avoid transient effects. The speed of every user is set to 1 m/s for simplicity. We use a truncated Pareto distribution as pause time distribution of which the minimum and maximum values are 30 seconds and 700 minutes, respectively.

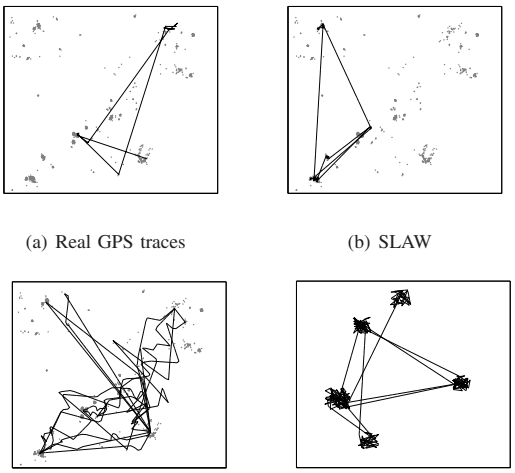
For simulation of various models, we use the following setup. In the Dartmouth model, the clusters are formed by the same method as in SLAW. It uses the same waypoints extracted from real traces [1] to build hot spots and also uses the transition probability obtained from the same traces. Note that since this model requires using these information, the simulation involving the Dartmouth model uses the same waypoint map that we obtained from the real traces in [1]. In the CMM model, the level of preferential attachment depends on the parameters such as the number of nodes and clustering exponent. We set the clustering exponent of the biggest hotspot to 0.5. In ORBIT, we vary the size of one side of hubs from 200 to 500 meters corresponding to the size of each simulation site while fixing the number of hubs. Each user selects the same number of hubs for daily travels as the number of hot spots chosen by individual walkers in SLAW.

In DTN routing protocol simulations, we generate one message bundle between each of randomly selected 100 source and destination pairs. All transmissions are assumed to be reliable and instantaneous when the communicating nodes are within a transmission range. To maximize the effect that mobility models have on routing performance, we assume that all nodes keep the entire history of past contacts with other nodes. All results are averaged over 40 runs.

B. Experimental Validation of SLAW

Figure 6 shows the sample traces of various mobility models. It is clearly visible that SLAW generates traces similar to real GPS traces. In the Dartmouth model [27], walkers visit every cluster with non-zero transition probability. In ORBIT [16], each user travels a fixed set of clusters (i.e., hubs) daily in a random order with a uniform probability. TLW [35] is a random model so it does not have common clusters for users. CMM [29] uses one popular cluster using preferential attachment but it also makes users visit every place in a given area. The traces of TLW and CMM are not shown to save space.

We now verify how well SLAW models the statistical features of human mobility. Since SLAW explicitly models $F2$ and $F4$, we focus only on truncated power-law flights and ICTs ($F1$ and $F3$). Figure 7 shows the flight distributions from various models and also from the measured GPS traces.



(a) Real GPS traces (b) SLAW

(c) Dartmouth (d) ORBIT

Fig. 6. Sample walk traces of various models.

SLAW is performed both on the waypoint maps generated synthetically by our fractal point generation technique and those extracted from real traces. SLAW produces very closely matching flight distributions (both synthetic and read maps) to that from the GPS traces. The Akaike test [10] tells whether the generated flight distributions fit power-law distributions (e.g., Pareto) or exponential distributions. Table II shows the result of the Akaike test [10] between Pareto and exponential distributions. In all cases, the flight distributions generated by SLAW are closer to truncated Pareto than expoential. All the other models except TLW do not produce power-law flights.

Figure 8 shows the distributions of ICTs among 50 nodes. Since we do not have any ICT traces corresponding to the GPS traces in [1], we cannot verify the realism of these ICT distributions. However, we can verify whether the ICT distributions follow a truncated power-law pattern. Table II shows the result of the Akaike test on the ICT distributions. It shows that the ICTs of SLAW, TLW and ORBIT fit better to power-law distributions while the ICTs of the other models fit better to exponential distributions. For the NYC traces, Dartmouth also shows Pareto fitting. SLAW tends to show many occurrences of short ICTs because of hot spots and regularity of visits. The ICTs of ORBIT shows significantly higher occurrences of very long ICTs. This is because in ORBIT, those with non-overlapping orbits do not meet at all while the others may also meet rarely because of randomness in picking the waypoints within their own orbits. The ICTs of CMM and Dartmouth have exponential distributions in most cases and tend to have much more occurrences of long ICTs than SLAW (note that the scales are log). This is because of their randomness in choosing the next clusters (or hot spots) to visit.

C. DTN Routing Performance

We test the following five DTN routing protocols. Random forwarding [37], Direct transmission [37], PRoPHET [30],

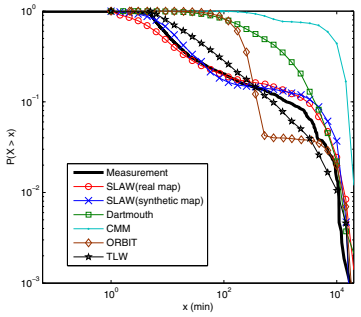


Fig. 7. Flight length distributions of synthetic traces from various models.

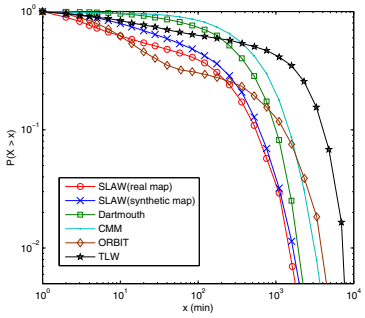


Fig. 8. ICT distributions of synthetic traces from various models

LET [19] and ECT. In *LET (last encounter time)*, a forwarding node of a message picks, as a next relay, the node with the most recent history of meeting the destination of the message among its current neighboring nodes. Each node updates its neighbor set at every minute. *ECT (expected contact time)* is a new metric we developed. It computes the expected time that a node meets the destination by subtracting the last encounter time from the *expected inter-contact time* which is computed by averaging the past inter-contact times with the destination.

We can categorize these protocols as *memoryless* and *memory-full* protocols. Random forwarding and direct transmission are memoryless as they do not use any past meeting history information. The other protocols are memory-full as they all use past contact information to predict the future probability of meeting the destination.

We find that the routing performance of TLW, RWP and CMM has almost the same patterns although their average delays. For brevity, we show the result of CMM only in Figure 9. In these models, both memoryless and memory-full protocols perform almost the same. This pattern happens because the mobility of nodes in these models is highly random so the prediction of memory-full protocols is not effective. The lack of performance differentiation among various types of protocols limits the usefulness of these models for mobile network simulation.

Dartmouth (Figure 10 (a)) shows results similar to CMM as the performance of various protocols except LET is not very distinguishable. The pattern can be explained as follows. The probability that Dartmouth nodes jump to any other hot spots

	NCSU		KAIST		NYC		Disney World	
	ICT	FL	ICT	FL	ICT	FL	ICT	FL
$\mathrm{SLAW}_{\mathrm{real}}$	Par	Par	Par	Par	Par	Par	Par	Par
$\mathrm{SLAW}_{\mathrm{syn}}$	Par	Par	Par	Par	Par	Par	Par	Par
Dartmouth	Exp	Exp	Exp	Exp	Par	Par	Exp	Exp
CMM	Exp	Exp	Exp	Exp	Exp	Par	Exp	Exp
ORBIT	Par	Exp	Par	Exp	Par	Exp	Par	Exp
TLW	Par	Par	Par	Par	Par	Par	Par	Par

TABLE II
THE RESULT OF THE AKAIKE TEST FOR THE MAXIMUM LIKELIHOOD ESTIMATION OF TRUNCATED PARETO DISTRIBUTIONS (DENOTED PAR) AND EXPONENTIAL DISTRIBUTION (DENOTED EXP) OVER FLIGHTS (DENOTED FL) AND ICTS EXTRACTED FROM SYNTHETICALLY GENERATED TRACES FROM VARIOUS MODELS WHOSE PARAMETERS ARE SET BASED ON REAL TRACES OBTAINED FROM FOUR DIFFERENT LOCATIONS (KAIST, NCSU, NYC AND DISNEY WORLD).

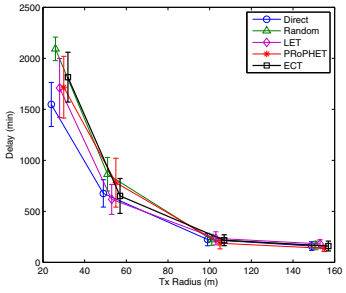


Fig. 9. Average routing delays of various protocols under CMM.

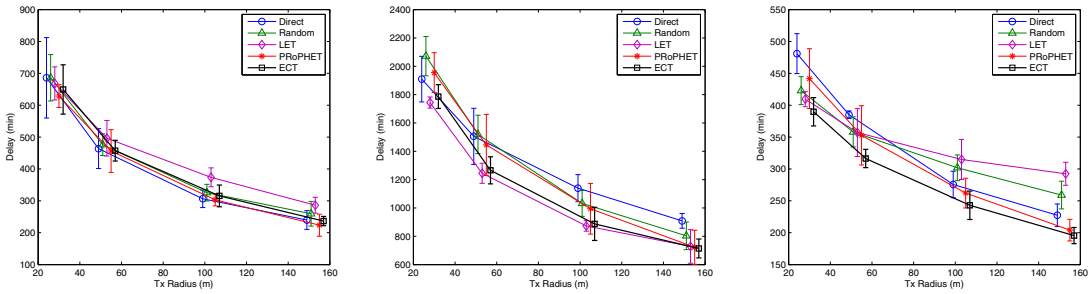
is determined by the transition probability; so any node can jump to any other hot spots as long as the transition probability is non-zero. This causes some randomness in visiting hot spots and likewise, ICT patterns similar to those in random mobility models such as CMM which are manifested in its exponentially distributed ICTs. However, because of higher transition probability to visit bigger hot spots, Darmouth exhibits much shorter delays than CMM and TLW. The low performance of LET in Darthmouth is because Dartmouth represents a hot spot as a single waypoint (i.e., a cluster reduces to one point) which has a side-effect of increasing pause times at a hot spot (since all the pause times for a hot spot are aggregated for all the points inside a hot spot). Thus, when a forwarding node meets a new node with a shorter LET than its LET, it is likely that the new node has just arrived to that hot spot. Thus the new node is more likely to stay in that hot spot much longer than the forwarding node, thus causing a longer delay to meet the destination next time. These features are an artifact of rather unusual and unrealistic setups of hot spots.

ORBIT (Figure 10 (b)) show a clear performance differentiation among different protocols. In ORBIT, nodes in non-overlapping "orbits" (i.e., they do not share a common hot spot) do not meet at all. The only way to deliver messages between two non-overlapping orbits is through the other nodes with overlapping orbits with the destination. This means that direct and random forwarding can perform really poorly. On the other hand, in ORBIT, memory-full protocols can perform much better. The performance difference between ECT and LET are relatively small compared to the difference between random forwarding and LET. This is because in ORBIT, the nodes with long LETs are likely to meet the destination fairly rarely due to only a small overlap in their orbits. Since each node in ORBIT moves like RWP among hot spots in its orbit, a small overlap results in a very long ICT and thus likely to have long LETs. Thus, the nodes with shorter LETs are likely to have more overlapping orbits with the orbit of the destination. Thus, choosing these nodes as relays leads to short routing delays. A similar argument is applicable for ECT because those with long LETs are likely to have long ECTs in ORBIT.

The mobility patterns of SLAW are quite different from those of ORBIT. The most salient feature is that SLAW has much shorter routing delays. SLAW has much more regularity in their trip patterns than any other models because it uses LATP for the selected set of waypoints and each node visits almost the same set of waypoints every day. In this type of scenarios, relaying to a node with short LETs can be detrimental because those nodes who just met the destination are likely to have long ECTs. That means that choosing, as relays, those nodes with short ECTs would always result in shorter routing delays because expected ICTs are very accurate because of the regularity in trip. The performance of PRoPHET is not as good as ECT because PRoPHET updates its probability only after meeting a destination. Thus, its behavior is slightly similar to that of LET. In SLAW, the power-law ICT distributions play a significant role. Due to the inspection paradox property of the renewal process and the power-law ICTs, when a node meets another node, it is more likely to meet a node with a long ICT. If it has a short ICTs, then ECT would perform as well as LET. But with long (predictable) ICTs, those nodes with short LETs are not likely to meet the destination for a long time. Thus ECT can perform much better than LET. The fact that ECT performs best indicates that the regularity of trip patterns is well represented in SLAW without loss of inherent statistical features such as power-law flight and ICT distributions.

V. CONCLUSION

In this paper, we present a new mobility model, called SLAW, that captures the effect of human mobility patterns found in real human mobility traces. We report many pieces of both analytical and empirical evidence that the movement of people can be expressed very well using gaps among fractal waypoints and present confirming data for the use of the least action principle in human trip planning. Based on this, we develop a simple heuristic algorithm called LATP that generates heavy-tail flights on top of fractal waypoints. Combining with heterogeneously bounded walkabout areas, we can successfully reproduce many statistical features important to the study of mobile network performance, especially, truncated power-law ICTs. Our routing performance study indicates that SLAW effectively expresses mobility patterns arising from peoples with some common interests or within a single community like students in the same university campus or people in theme parks where people tend to share common gathering places. We find that LATP over heterogeneously bounded areas



(a) Dartmouth (b) ORBIT (c) SLAW

Fig. 10. Average routing delays of various protocols under Dartmouth, ORBIT and SLAW.

realistically expresses some periodicity in the daily mobility of humans. This feature makes many memory-full routing protocols such as utility-based routing very effective.

REFERENCES

[1] http://netsrv.csc.ncsu.edu/levy-mobility, 2008.

[2] F. Bai, N. Sadagopan, and A. Helmy. Important: a framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks. In *Proceedings of INFOCOM 2003*, San Francisco, CA, April 2003.

[3] A. Balasubramanian, B. N. Levine, and A. Venkataramani. Dtn routing as a resource allocation problem. In *Proceedings of SIGCOMM 2007*, Kyoto, Japan, August 2007.

[4] J. Banerjee, M. K. Verma, S. Manna, and S. Ghosh. Self-organised criticality and 1/f noise in single-channel current of voltage-dependent anion channel. *Europhysics Letters*, 73(3):457–463, February 2006.

[5] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[6] A.-L. Barabasi and E. Bonabeau. Scale-free networks. *Scientific American*, 288:50–59, 2003.

[7] V. Borrel, M. D. de Amorim, and S. Fdida. A preferential attachment gathering mobility model. *IEEE Communications Letters*, 9:900–902, 2005.

[8] V. Borrel, F. Legendre, M. D. de Amorim, and S. Fdida. Simps: Using sociology for personal mobility. *submitted to IEEE Transactions on Networking*, 2007.

[9] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439:462–465, January 2006.

[10] K. P. Burnham and D. R. Anderson. Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods and Research*, 33(2):261–304, November 2004.

[11] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on the design of opportunistic forwarding algorithms. In *Proceedings of INFOCOM 2006*, Barcelona, Spain, April 2006.

[12] D. R. Cox. Long-range dependence: A review. In *Statistics : An Appraisal*, pages 55–74. Ames : Iowa State University Press, 1984.

[13] P. L. M. de Maupertuis. Accord des differentes lois de la nature qui avaient jusqu'ici paru incompatibles. *Memoires de l'Academie des Sciences*, page 417, 1744.

[14] F. Ekman, A. Keränen, J. Karvo, and J. Ott. Working day movement model. In *MobilityModels '08: Proceeding of the 1st ACM SIGMOBILE workshop on Mobility models*, pages 33–40, 2008.

[15] J. Gaite. Zipf's law for fractal voids and a new void-finder. *The European Physical Journal B: Condensed Matter and Complex Systems*, 47(11):93–98, 2005.

[16] J. Ghosh, S. J. Philipb, and C. Qiao. Sociological orbit aware location approximation and routing (solar) in manet. *Ad Hoc Networks*, 5:189–209, 2007.

[17] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. *Nature*, 453:779–782, June 2008.

[18] J. Gordon. Pareto process as a model of self-similar packet traffic. In *Proceedings of GLOBECOM 1995*, pages 2232–2236, Singapore, November 1995.

[19] H.Dubois-Ferriere, M.Grossglauser, and M.Vetterli. Age matters: Efficient route discovery in mobile ad hoc networks using encounter ages. In *Proceedings of MobiHoc 2003*, Annapolis, MD, June 2003.

[20] S. Hong, I. Rhee, S. J. Kim, K. Lee, and S. Chong. Routing performance analysis of human-driven delay tolerant networks using the truncated levy walk model. In *Proceeding of the 1st ACM SIGMOBILE workshop on Mobility models*, Hong Kong, May 2008.

[21] X. Hong, M. Gerla, G. Pei, and C. Chiang. A group mobility model for ad hoc wireless networks. In *Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems (MSWiM 1999)*, pages 53–60, Seattle, WA, August 1999.

[22] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy. Modeling time-variant user mobility in wireless mobile networks. In *Proceedings of INFOCOM 2007*, Anchorage, AL, May 2007.

[23] E. Hyytia, P. Lassila, and J. Virtamo. A markovian waypoint mobility model with application to hotspot modeling. In *Proceedings of ICC 2006*, Istanbul, Turkey, June 2006.

[24] A. Jardosh, E. Belding-Royer, K. Almeroth, and S. Suri. Towards realistic mobility models for mobile ad hoc networks. In *Proceedings of MobiCom 2003*, pages 217–229, San Diego, CA, September 2003.

[25] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic. Power law and exponential decay of inter contact times between mobile devices. In *Proceedings of MobiCom 2007*, Montreal, Canada, September 2007.

[26] M. Kim and D. Kotz. Periodic properties of user mobility and access-point popularity. *Personal Ubiquitous Computing*, 11(6):465–479, 2007.

[27] M. Kim, D. Kotz, and S. Kim. Extracting a mobility model from real user traces. In *Proceedings of INFOCOM 2006*, Barcelona, Spain, April 2006.

[28] B. Liang and Z. J. Haas. Predictive distance-based mobility management for multidimensional pcs networks. *IEEE/ACM Transactions on Networking*, 11(5):718–732, October 2003.

[29] S. Lim, C. Yu, and C. R. Das. Clustered mobility model for scale-free wireless networks. In *Proceedings of the 31st IEEE Conference on Local Computer Networks (LCN 2006)*, Tampa, FL, November 2006.

[30] A. Lindgren, A. Doria, and O. Schelen. Probabilistic routing in intermittently connected networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 7(3):19–20, July 2003.

[31] B. B. Mandelbrot. *The Fractal Geometry of Nature*. W.H. Freeman and Company, New York, 1977.

[32] I. Norros, P. Mannersalo, and J. L. Wang. Simulation of fractional brownian motion with conditionalized random midpoint displacement. In *Advances in Performance Analysis*, pages 77–101, 1999.

[33] I. Norros and J. Virtamo. Handbook of fbm formulae, September 1996.

[34] I. Rhee, K. Lee, S. Hong, S. J. Kim, and S. Chong. Demystifying the levy-walk nature of human walks. In *Techical Report, NCSU, http://netsrv.csc.ncsu.edu/export/Demystifying_Levy_Walk_Patterns.pdf*, 2008.

[35] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong. On the levy walk nature of human mobility. In *Proceedings of INFOCOM 2008*, Phoenix, AZ, April 2008.

[36] H. G. Schuster. *Deterministic Chaos*. VCH Publisher, New York, 1989.

[37] T. Spyropoulos, K. Psounis, and C. S. Raghavendra. Efficient routing in intermittently connected mobile networks: The single-copy case. *IEEE/ACM Transactions on Networking*, 16(1):63–76, 2008.