Smoothly Truncated Levy Walks: Toward a Realistic Mobility Model

Lijuan Cao
Johnson C. Smith University
Charlotte, NC, USA
Email: lcao@jcsu.edu

Michael Grabchak
University of North Carolina at Charlotte
Charlotte, NC, USA
Email: mgrabcha@uncc.edu

***Abstract*—Mobility models are crucial for the simulation and evaluation of protocols for multihop wireless networks. However, most commonly used mobility models do not reflect the way humans actually move. This significantly affects the reliability of simulation results. In this paper we introduce a new mobility model called the Smoothly Truncated Levy Walk (STLW), which is more realistic than most standard models that appear in the literature. Its main innovations are as follows. First, to take into account dependencies in the direction of motion, it models changes in the direction instead of the standard approach, which directly models the direction and ignores these dependencies. Second, it uses realistic models for pause times, flight lengths, and changes in direction. In particular, it uses tempered stable distributions to model pause times and flight lengths and the beta distribution to model changes in direction. We justify the use of these distributions from both a theoretical and an empirical perspective. In particular, we perform a trace-based validation on several real-world traces from various scenarios. Validation results show that this model is very flexible and can be used to model human movements in a variety of situations.**

I. Introduction

Multihop wireless networks (e.g. Mobile Ad Hoc Networks (MANET), Vehicular Ad Hoc Networks (VANET), and Delay Tolerant Networks (DTN)) have drawn a great deal of research effort over the past decade. Unlike more standard networks (e.g. cellular networks or WLAN) they do not have an established infrastructure or a centralized controller. Instead, they rely on the mobile nodes themselves (i.e. mobile devices carried by humans) to maintain network connectivity and relay messages to each other. This infrastructureless structure makes network deployment easy, but it has several unique features, such as peer-to-peer multihop communications and a dynamic topology due to node movements, which make routing protocol design a challenging task.

Simulation is perhaps the most widely used tool in the development, evaluation, and refinement of multihop wireless network protocols. It has the advantages of scalability, reproducibility, and time and cost efficiency. In order to run simulations one needs a mobility model to generate the random movements of humans. Research has shown that different mobility models can significantly impact the performance of MANET routing protocols, including the packet delivery ratio, the control overhead, and the packet delay, see [2]. Therefore, in order to accurately evaluate performance, it is important to use a realistic mobility model that can accurately reflect the movement of humans.

A large number of mobility models have been proposed in the literature. Building on these, we define a new mobility model called the Smoothly Truncated Lévy Walk (STLW). This model uses tempered stable distributions to model flight lengths and pause times and a beta distribution to model changes in direction. These are flexible models, which do not appear to have been used in this context before. Moreover, the idea of modeling changes in the direction travelled appears to be new. Most models assume that the direction travelled is independent of the previous direction, but this assumption ignores important dependence structures seen in the data.

In Section II we discuss related work, in Section III we present our model, in Section IV we give a theoretical motivation for this model, and in Section V we fit our model to real-world data and provide an empirical justification for it. The real-world data consists of four traces collected in [19], which represent mobility characteristics of wireless network users in various scenarios. In Section VI we use simulation to study the topological properties of this model. Finally, we conclude in Section VII by discussing several directions for future work.

II. Related Work

A number of different mobility models have appeared in the literature. Many of these use historical or geographic information to generate mobility traces, see e.g. [12], [14], [17], [11]. While this is useful for developing protocols to be used in specific scenarios, we focus on general mobility models that are not linked to particular situations.

The Random Way Point (RWP) model [13] is, perhaps, the best known and most widely used mobility model. It generates traces (i.e. paths) as follows. Each node starts at a randomly selected location in the simulation area. It stays there for a fixed amount of time, which is called the pause time. After this it chooses a destination randomly and moves toward this destination at a speed randomly chosen from between some predetermined minimum and maximum values. When the node reaches its destination, it stays there for the fixed pause time and then the process repeats. Although a constant pause time and the way destinations are chosen are not realistic, nevertheless RWP provides the basic ingredients that a mobility model should have.

To flesh this out further, a number of modifications to RWP have been proposed. One of the best known is the Random Direction (RD) model introduced in [20]. Here each node chooses a direction uniformly at random and then travels in this direction until it reaches the boundary of the simulation area. Then, after a fixed pause time, it chooses a new direction and the procedure repeats. Since, in practice, a person may stop or change direction before reaching a boundary, [20] also introduced the Modified Random Direction (MRD) model. Here the node chooses a random direction and then moves in this direction for a random distance. This distance, which is called the flight length, is chosen uniformly at random.

In [19], after a thorough analysis of several real-world traces, a modification of MRD was proposed. In this modification the pause times are no longer fixed. Instead they are random and the distributions of both the pause times and the flight lengths are much more realistic. Specifically, it is assumed that both of these are simulated as follows. A proposed pause time and a proposed flight length are simulated from (potentially different) symmetric stable distributions. If the values are negative then they are ignored and new values are sampled. Further, there are truncation parameters $\tau_p, \tau_f > 0$. If a simulated pause time exceeds τ_p then the observation is ignored and a new one is simulated. A similar procedure is preformed to ensure that the flight lengths do not exceed τ_f. This model is called the Truncated Lévy Walk (TLW).

Although it is more realistic than the others, the TLW model, nevertheless, has several limitations. These include the distributions that are used and the way in which new directions are chosen. We will discuss these limitations and ways to improve upon them in the next section. Before proceeding, we mention that modifications of RWP in somewhat different directions have also been proposed in [1] and [3].

III. Smoothly Truncated Levy Walks

In this section we describe our mobility model. It is inspired by the TLW model introduced in [19], but it differs from that model in the following important ways:

- We allow for the possibility that a person may change direction without pausing.
- Instead of directly modeling the direction in which a person moves, we model the change in this direction. For this we use a beta distribution.
- We use tempered stable distributions to model pause times and flight lengths.

The first of these modifications is self explanatory. The second stems from the fact that people tend to travel with a purpose. While there may be various geographic obstacles forcing them to deviate from a straight line, on the whole they tend to go toward a particular destination. This creates dependencies in the direction travelled, which are taken into account by modeling changes in the direction.

Tempered stable distributions are similar to the ones used in the TLW model, in that they are very similar to stable distributions in some central region, but their tails have been modified to ensure that the values do not get too big. The main difference is the way in which this modification is made. In the TLW model the tails are modified through a "hard" truncation, which relies on arbitrary truncation parameters that are difficult to justify from a practical perspective. After all there do not exist reasonable values such that one can say with certainty that a person may wait (or move) less than this value but not more. On the other hand, tempered stable distributions have a "smooth" truncation, where their tails are made to, eventually, decay exponentially fast. For this reason we call our models Smoothly Truncated Lévy Walks (STLW).

Further, tempered stable distributions have nice analytic properties, which allow us to, among other things, give a theoretical justification for their use, see Section IV. Below, we introduce the probability models that we will use and present the algorithm for STLW. The probability models are summarized in Table 1.

A. Distributions

(Model of changes in direction.) To model changes in direction we use the beta distribution, which is one of the best known and most heavily used distributions on a finite interval. The beta distribution on $(0, \pi)$ has a density given by

$$f(x) = \begin{cases} \frac{\pi^{1-\alpha-\beta}}{B(\alpha,\beta)} x^{\alpha-1}(\pi - x)^{\beta-1}, & 0 < x < \pi \\ 0, & \text{otherwise} \end{cases},$$

where $B(\cdot, \cdot)$ is the beta function and $\alpha, \beta > 0$ are parameters. We denote this distribution by $\mathrm{beta}(\alpha, \beta)$. The parameter α describes what happens near zero and the parameter β describes what happens near π. In particular, if $\alpha \in (0, 1)$ then the density is unbounded near 0 creating a mode there. A similar fact holds near π when $\beta \in (0, 1)$. It can be readily checked that, if X has distribution $\mathrm{beta}(\alpha, \beta)$ then $\mathrm{E}[X] = \pi\frac{\alpha}{\alpha+\beta}$ and $\mathrm{Var}(X) = \pi^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$. An important special case is when $\alpha = \beta = 1$. Here the model reduces to the uniform distribution on $(0, \pi)$.

(Model of flight lengths and pause times.) To model flight lengths and pause times we use the tempered stable distribution. A tempered stable distribution on $[0, \infty)$ has three parameters, $\alpha \in (0, 1)$, $\beta > 0$, and $\ell > 0$. Its moment generating function is given by

$$\mathrm{E}[e^{zX}] = \exp\left\{\frac{\beta}{\ell^\alpha}\left[1 - (1 - \ell z)^\alpha\right]\right\}, \quad z \le 1/\ell.$$

We denote this distribution by $TS(\alpha, \beta, \ell)$. It is easy to check that $\mathrm{E}[X] = \beta\alpha\ell^{1-\alpha}$ and $\mathrm{Var}(X) = \beta\ell^{2-\alpha}\alpha(1 - \alpha)$. These distributions were first introduced in [21], see [16] for a more recent review. They form an important subclass of a more general class of tempered stable models, see [7] and the references therein.

While the densities of these distributions are not available in a closed form, very good numerical estimates are available in the Tweedie library [4] for the statistical software package R. This library also has methods for simulating tempered stable random variables and evaluating their cumulative distribution functions. To get a better understanding of the role played by the parameters, we give the following asymptotic result. If g is the density of a $TS(\alpha, \beta, \ell)$ distribution, then

$$g(x) \sim Ce^{-x/\ell}x^{-1-\alpha} \text{ as } x \to \infty, \tag{1}$$

where $C = \frac{\beta\alpha}{\Gamma(1-\alpha)}e^{\beta/\ell^\alpha}$ and $\Gamma(\cdot)$ is the gamma function.

TABLE I. PROBABILITY MODELS.

Model	Notation	Parameters
Stable	$S(\alpha,\beta)$	$\alpha \in (0,1)$, $\beta > 0$
Tempered Stable	$TS(\alpha,\beta,\ell)$	$\alpha \in (0,1)$, $\beta > 0$, $\ell > 0$
Beta	$\mathrm{beta}(\alpha,\beta)$	$\alpha > 0$, $\beta > 0$

The name "tempered stable distribution" is explained by the fact that these distributions have densities that are very similar to those of stable distributions in some central region, but they have lighter, i.e. tempered, tails. To formalize this, first recall that stable distributions on $[0,\infty)$ are determined by

$$\mathrm{E}[e^{zY}] = e^{-\beta|z|^{\alpha}}, \quad z \le 0,$$

where $\alpha \in (0,1)$ and $\beta > 0$ are parameters. We denote this distribution by $S(\alpha,\beta)$.

Although stable distributions have important theoretical interpretations (see Theorem 1 in Section IV) and have been found to be very useful in a variety of applications, they have extremely heavy tails, and, in fact, both their means and their variances are infinite. This is not realistic for most applications since, in practice, there are all kinds of real-world obstacles limiting the size of random phenomena. For this reason it is useful to consider modifications of these distributions, which have lighter tails. One way to do this is as follows. Let $f(x)$ be the density of a $S(\alpha,\beta)$ distribution and let

$$g(x) = Ke^{-x/\ell} f(x),$$

where $K = e^{\beta/\ell^{\alpha}}$. This distribution has all moments finite, and, in fact, it is the density of a $TS(\alpha,\beta,\ell)$ distribution. It is clear that if $\ell > 0$ is very large then this density will be indistinguishable from $f(x)$ for small, medium, and even fairly large values of x, but for very large values of x the tails decay exponentially fast.

B. Algorithm

We now describe the implementation of our mobility model. We begin with a node placed on the simulation area. It can be placed at a random or a prechosen location. Since there is no reason to believe, a priori, that a person is more likely to move in any particular direction, we choose the initial direction uniformly at random. After this, we choose a flight length from a $TS(\alpha_f,\beta_f,\ell_f)$ distribution and we choose a speed uniformly at random from between some preset minimum and maximum values. Once we reach the destination, we wait a random amount of time. With probability p this time is zero and with probability $1-p$ we sample this time from a $TS(\alpha_p,\beta_p,\ell_p)$ distribution. After this time has elapsed, we sample a direction change θ from a $\mathrm{beta}(\alpha_d,\beta_d)$ distribution and a flight length X from a $TS(\alpha_f,\beta_f,\ell_f)$ distribution. With probability .5 we turn θ radians to the left or with probability .5 we turn θ radians to the right. Then we choose a speed as before and travel a distance of X in the new direction at this speed. We then iterate the procedure. See Algorithm 1 for more details.

In practice, one generally wants the nodes to stay within a fixed and finite simulation area. To ensure this, we suggests the following modification to the algorithm. When a node reaches the boundary it should bounce off such that the angle of incidence equals the angle of reflection.

Parameters: For pause time: $p \in (0,1)$, $\alpha_p \in (0,1)$, $\beta_p > 0$, $\ell_p > 0$; For flight length: $\alpha_f \in (0,1)$, $\beta_f > 0$, $\ell_f > 0$; For direction change: $\alpha_d > 0$, $\beta_d > 0$; For speed: $\mathrm{maxspeed} > \mathrm{minspeed} > 0$.
Begin:
1. Choose a starting location on the simulation area.
2. Simulate a direction ϕ uniformly at random and simulate a flight length X from $TS(\alpha_f,\beta_f,\ell_f)$.
3. Choose a speed uniformly between $\mathrm{minspeed}$ and $\mathrm{maxspeed}$ and move in direction ϕ a distance of X at this speed.
while *Arrive at destination* **do**
  4. With probability p simulate a random time from $TS(\alpha_p,\beta_p,\ell_p)$ and wait this amount of time. Otherwise continue to step 5.
  5. Simulate θ from $\mathrm{beta}(\alpha_d,\beta_d)$ and X from $TS(\alpha_f,\beta_f,\ell_f)$.
  6. With probability .5 turn θ radians left otherwise turn θ radians right.
  7. Choose a speed uniformly between $\mathrm{minspeed}$ and $\mathrm{maxspeed}$ and move forward a distance of X at this speed.
end

Algorithm 1: STLW mobility model.

IV. THEORETICAL JUSTIFICATION

In this section we give a theoretical justification for using the tempered stable distribution to model flight lengths. To do this we begin by describing the motion of a person, who we will refer to as the walker. First, assume that we observe the location of the walker at fixed time increments Δt. Let $\{X_n : n = 0,1,\dots\}$ be a discrete time stochastic process on $\mathbb{R}^2$ such that X_n is the location of the walker at time $n\Delta t$. Let $Z_n = X_n - X_{n-1}$ be the increment process. If the walker did not stay in the same place during the ith time interval we can write $Z_i = \frac{Z_i}{|Z_i|}|Z_i|$, where $|Z_i|$ represents the magnitude of the displacement and $\frac{Z_i}{|Z_i|}$ represents the direction of travel.

A common model for $|Z_i|$ is called a Lévy walk. Here it is assumed that the magnitudes of displacement $|Z_i|$ are iid random variables having a Pareto distribution, i.e. having a density given by

$$f(x) = \begin{cases} \alpha\delta^{\alpha}x^{-\alpha-1} & x > \delta \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

where $\alpha \in (0,1)$ and $\delta > 0$ are parameters. One can allow α to be any positive number, but most empirical data on human mobility suggests $\alpha \in (0,1)$, see [6], [19], [18], and the references therein. In practice, however, the movement of humans is more complicated and the Pareto distribution is only valid for large values of x. For this reason we assume only that for some $c > 0$ the density satisfies

$$f(x) \sim cx^{-\alpha-1} \text{ as } x \to \infty. \tag{3}$$

It turns out that the actual distribution of $|Z_i|$ does not matter if we model the lengths of entire flights at once. A flight is a part of the walk where the walker keeps going in the same direction. Assume, for instance, that after the walker starts walking at time 0, the first time that she stops

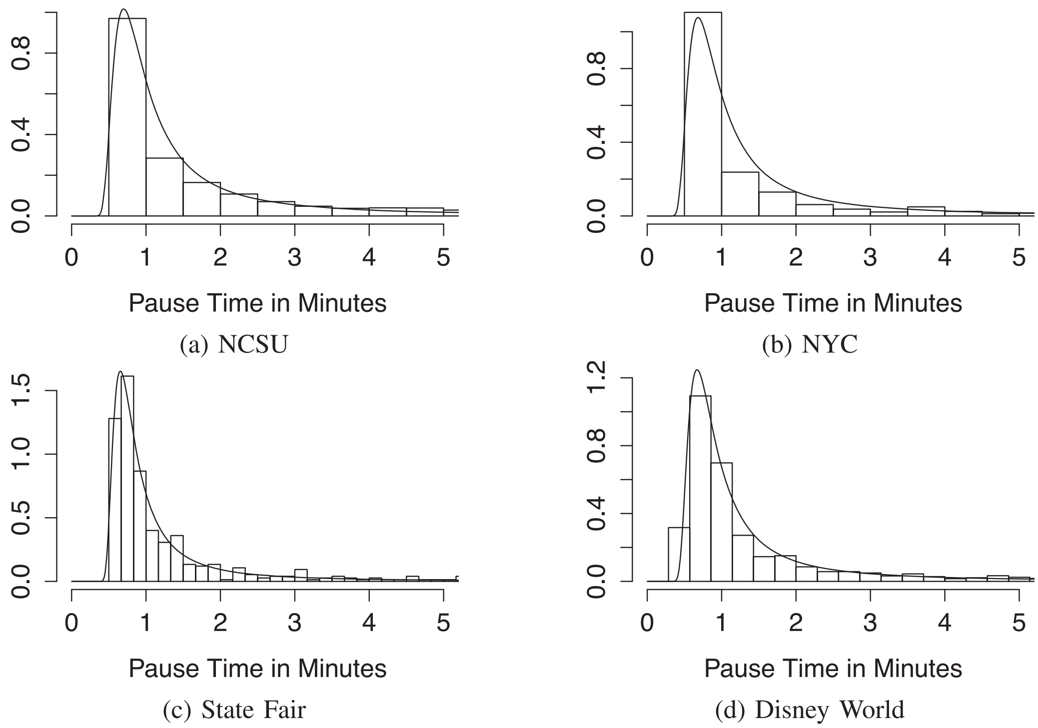


Fig. 1. The histograms (normalized to sum to one) represent the empirical distributions of the pause times. Overlaid are plots of the densities of the estimated tempered stable distributions.

or changes direction is at time $N\Delta t$. This means that $|X_N| = |Z_1 + Z_2 + \cdots + Z_N| = \sum_{i=1}^N |Z_i|$, where the second equality follows because all of the steps are in the same direction. Since a person is likely to walk in the same direction for a relatively long time, N is likely to be quite large and we can approximate the distribution of $|X_N|$ by its asymptotic distribution. This asymptotic distribution is described in the theorem below, which is a version of the central limit theorem for infinite variance distributions, see [5] for details.

Theorem 1. *If $\alpha \in (0,1)$ then*

$$n^{-1/\alpha} \sum_{i=1}^{n} |Z_i| \xrightarrow{d} Y,$$

where Y has a $S(\alpha, \beta)$ distribution and $\beta > 0$ is a constant depending of the distribution of $|Z_i|$.

This discussion suggests that flight lengths should be well modeled by $S(\alpha, \beta)$ distributions. However, the tails of these distributions are too heavy. In practice, there are various geographic and physical limitations that prevent flight sizes from getting too big. In fact, empirical data suggests that (3) holds for large, but not too large values of x, see [6], [19], [18], and the references therein. This has led to the development of tempered Lévy walks.

A tempered Lévy walk assumes that the density satisfies

$$f(x) \sim ce^{-x/\ell} x^{-\alpha-1} \text{ as } x \to \infty, \tag{4}$$

for $\alpha \in (0,1)$ and $\ell > 0$. If ℓ is very big, this means that for medium and somewhat large values of x $f(x) \approx cx^{-\alpha-1}$, but for very large values of x, we start to feel the exponential function and the tails ultimately decay exponentially fast. We will show that, in this case, the sum $|X_N| = \sum_{i=1}^N |Z_i|$ is well approximated by a tempered stable distribution. Although we can no longer use the central limit theorem, we can use the following result from [9] (for a preliminary version see [8]). First, we introduce the notation $Z_i(\ell) = Z_i$ to emphasize the dependence of the distribution of Z_i on the parameter ℓ.

Theorem 2. *If $\ell \to \infty$ such that $n^{-1/\alpha}\ell \to \ell' \in (0, \infty)$ then*

$$n^{-1/\alpha} \sum_{i=1}^{N} |Z_i(\ell)| \xrightarrow{d} Y \text{ as } n, \ell \to \infty,$$

where Y has a $TS(\alpha, \beta, \ell')$ distribution and β is a parameter depending on the distribution of $|Z_i(\ell)|$.

In practice, of course, the parameter ℓ is not actually approaching infinity. Instead it is some fixed but (very) large constant. In this case we can interpret the theorem in the following way. If n is large, but $n^{-1/\alpha}\ell$ is medium sized, we can approximate the distribution of the sum by a tempered stable distribution. A different justification for approximating the sum by a tempered stable distribution is given in [10]. There the approximation follows by showing that the distribution of the sum is close to that of a tempered stable distribution in a certain metric on the space of probability distributions. This discussion suggests that tempered stable distributions should be very good models for flight lengths.

V. Trace analysis

In this section we validate our models by fitting them to real world data. The data consists of human mobility traces from four different sites collected by [19]. These sites are a state fair in North Carolina, Disney World in Florida, the campus

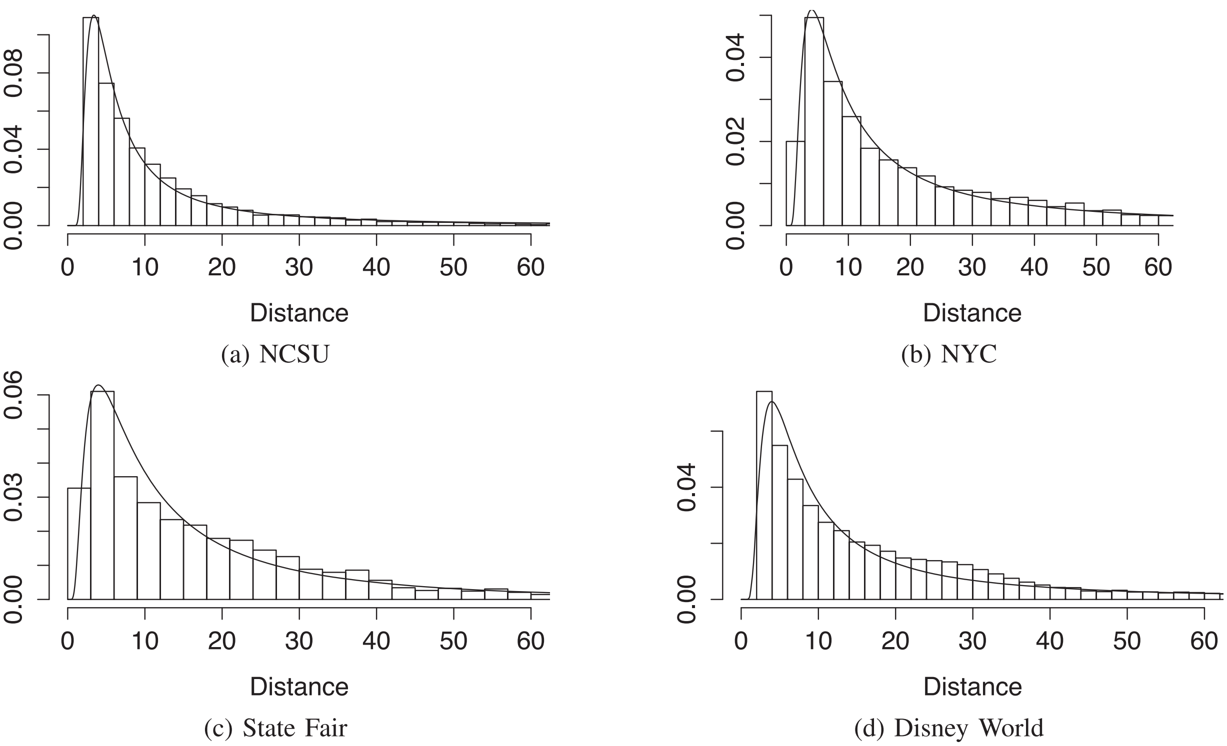


Fig. 2. The histograms (normalized to sum to one) represent the empirical distributions of the flight lengths. Overlaid are plots of the densities of the estimated tempered stable distributions.

of North Carolina State University (NCSU), and New York City (NYC). They range from small to very large geographic areas, and the traces collected from them represent human movements using different modes of travel. We obtained the trace data from CRAWDAD [15]. All of the data had been collected using Garmin GPS 60CSx handheld receivers, which took samples every 10 seconds. Each position was then recomputed as an average of three samples over the 30 second period to account for GPS errors. In processing the data, we extracted information about flight lengths, pause times, and direction changes as follows. We consider a pause to be any period when the distance travelled between two consecutively sampled positions (i.e. during a 30 second period) is less than a predefined threshold value. Consecutive pause periods are considered to be part of the same pause. We define the start of a flight to be one of two situations. Either a) this is the first time we move more than the threshold value after a pause or b) the change in direction between two consecutive periods is more than a predefined threshold value. If we are not in a pause and we are not starting a new flight, then we assume that the previous flight is continuing. With these definitions, we extract the following information from the trace data:

- Flight length: the distance travelled during one flight.
- Pause time: the time elapsed between two consecutive flights.
- Direction change: the difference in direction between two consecutive flights.

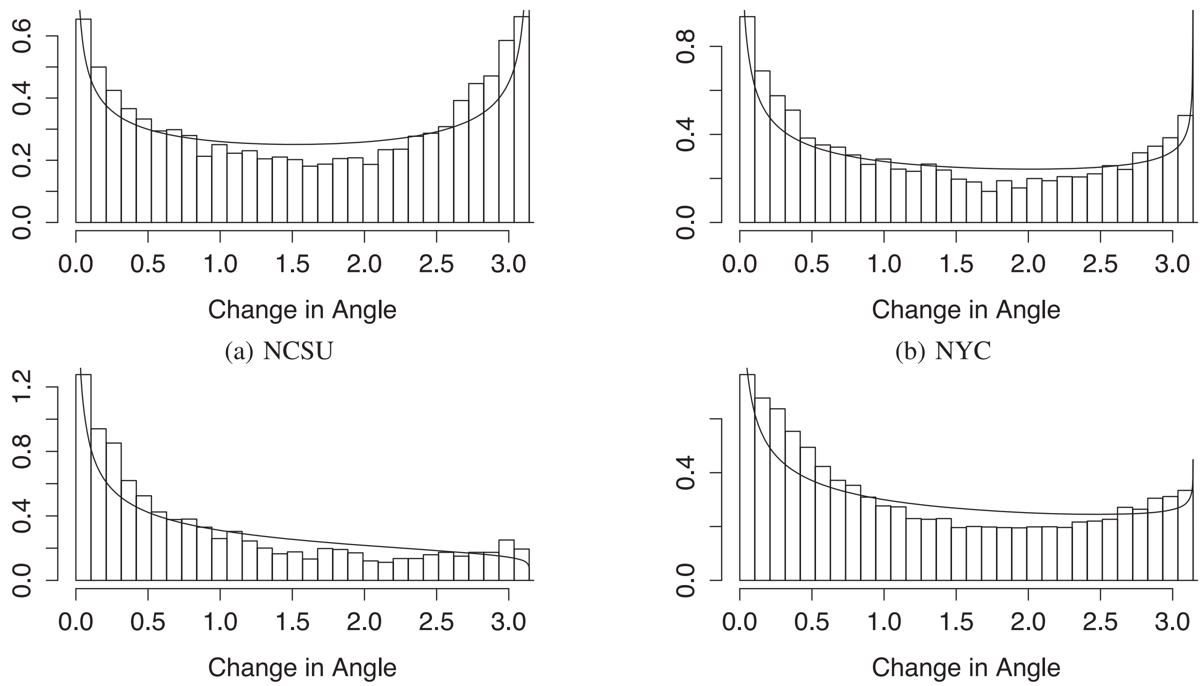
For each of the four scenarios, we used the data to estimate the parameters of our model. These parameters were estimated using the method of maximum likelihood. This method requires initial estimates, which were found using the method of moments. The density of the tempered stable distribution was evaluated using the Tweedie library [4] for the statistical software package R. In the following subsections we present and discuss the results of this analysis.

A. Flight lengths and pause times

The data sets that we use were collected and first analyzed in [19]. There it was shown that the tails of the distributions of flight lengths and pause times decay like a polynomial up to a point, but, eventually, the decay becomes exponentially fast. This is exactly the behavior exhibited by tempered stable distributions, see equation (1). We will now see how well these models fit the data.

Figure 1 shows histograms of the pause times for each of the four scenarios. Here (and throughout) the histograms have been normalized to sum to one. Overlaid are plots of the densities of the estimated tempered stable distributions. We see that this model fits the data very well. It does so both in the mode and in the tails.

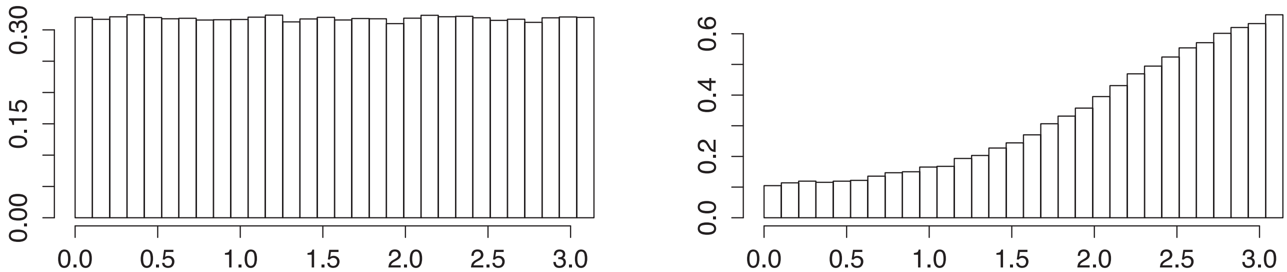
Figure 2 shows our analysis of the flight lengths. For each trace we give a histogram of the data overlaid with the density of the estimated tempered stable distribution. We see that the model does a very good job fitting the data in the case of NCSU and NYC. These are situations with a large geographic area, where locations of interest may be quite far apart. The fit is slightly worse in the case of the state fair and Disney World. In the plot for Disney World we see that, although the estimated tempered stable density does a good job modeling the mode and the tail, the histogram has quite a bit of mass above the estimated density in the central portion. This suggests that the tempered stable distribution under-estimates



(a) NCSU (b) NYC

(c) State Fair (d) Disney World

Fig. 3. The histograms (normalized to sum to one) represent the empirical distributions of the changes in direction. Overlaid are plots of the densities of the estimated beta distributions.



(a) Change in direction for MRD and TLW (b) Change in direction for RWP

Fig. 4. Simulated distributions of direction changes in common mobility models.

the chance of a "medium" sized flight. This may be caused by the fact that, in Disney World, there are many attractions relatively close to each other, and a visitor is more likely to go to an attraction that is not very far away. A similar situation is seen in the state fair data. Never-the-less, the tempered stable distribution captures the main trends seen in these cases.

This discussion suggests that when we want to model specific scenarios (i.e. if we are particularly interested in amusement parks) it is useful to take into account the specific features of those situations. However, when we are interested in a general situation, tempered stable distributions provide a flexible class of models that cover the basic trends seen in the data and have an intuitive theoretical interpretation.

B. Direction Change

In this section we analyze the direction change data. Note that we only consider changes in direction when a new flight begins. Figure 3 shows histograms of the changes in direction for each of the four scenarios. Overlaid are plots of the densities of the estimated beta distributions. From the histograms we see that the probability of making small changes in direction is quite high. Intuitively, this makes sense since people tend to travel with a purpose. While there may be various geographic obstacles forcing them to deviate from a straight line, on the whole they tend to go toward a particular destination. In the NCSU and NYC data we see that the chance of turning around close to π radians ($180°$) is also relatively high. Intuitively, this may be because once a person is finished at a destination she is more likely to turn around and go back. This kind of behavior is much less prominent in the state fair and Disney World data. This may be because at these locations there are many attractions near each other and hence there is less need to go back.

We now turn our attention to the estimated beta distributions. We see that, although they are not perfect, they capture the main trends seen in all four datasets. In particular, they capture both the mode near 0 and the smaller one near π. The only exception is the state fair data, where it is less clear if a mode near π truly exists. From the structure of the data it is

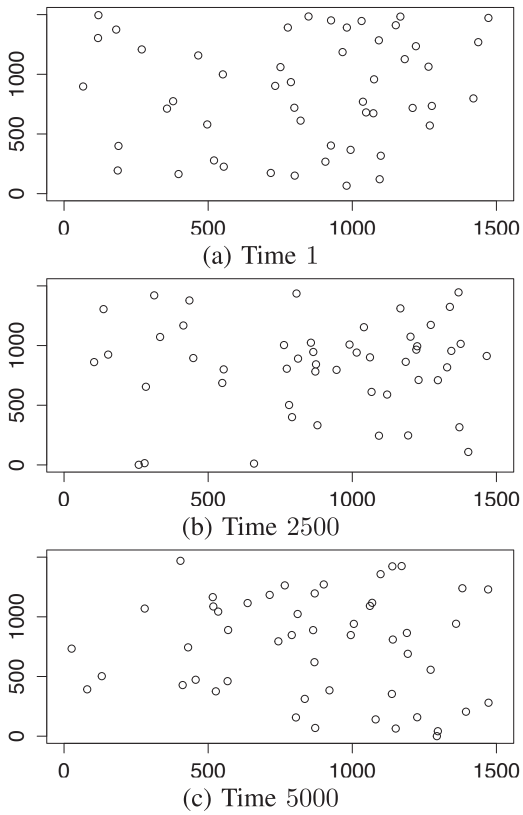


Fig. 5. Locations of the nodes at three time points. They are evenly distributed with no parts of the simulation area having a large cluster of nodes.

clear that, for the purposes of simulation, we should choose our parameters to satisfy $\alpha \in (0, 1)$ and $\alpha < \beta$. This will ensure that there is a mode at 0 and that the mode near π (if it exists) is smaller than the one near 0. If we want to guarantee the existence of a mode near π we should choose $\beta \in (\alpha, 1)$.

Having shown that the beta distribution captures the general trends observed in the data, we now show that the standard models do not. To do this we simulated a trace with 5000 flights from each of two common models. We used this to estimate the distribution of changes in direction. In Figure 4a) we give the histogram of changes in direction for the MRD model (note that directions are chosen in the same way in the TLW model). In this case, since the directions are chosen uniformly at random, it is not surprising that the direction changes are uniform as well. In Figure 4b) we give the histogram of changes in direction for the RWP model based on a 500×500 simulation area. In this case changes in direction tend to be very large. This means that a person following the RWP model is much more likely to go backward than forward. This seems to be caused by the fact that, in the RWP model, a person is more likely to move toward the center of the simulation area than anywhere else. Since, in the case of RWP, the distribution of changes in direction is sensitive to the dimensions of the simulation area, we tried a variety of different dimensions. However, the general trends were similar in all cases. When we compare the histograms of direction changes in these two models with the histograms of real-world data given in Figure 3 we see that they are very different, and hence these models do not model the data well. On the other hand, the beta distributions capture exactly the

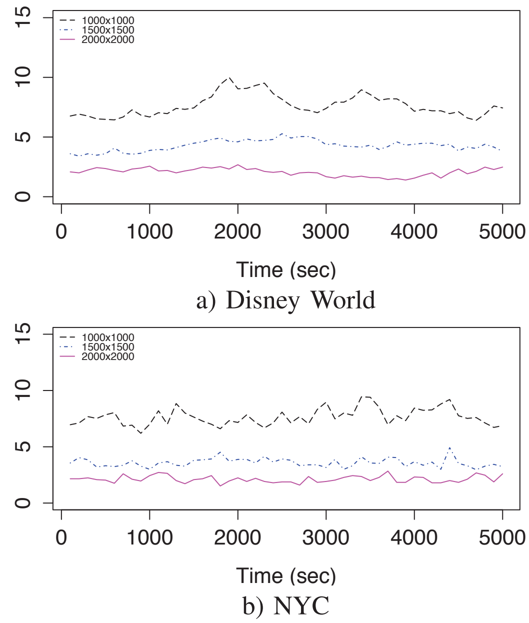


Fig. 6. Average number of neighbors in three different simulation areas in the Disney World and NYC scenarios.

kind of behavior observed in the data.

VI. Simulations

When developing efficient routing protocols for multihop wireless networks it is important to understand the topological structure of the nodes and how this structure evolves over time. To study this we conduct several sets of simulations. For these simulations, mobility traces are generated based on the parameters estimated in two scenarios: Disney World and NYC. For both scenarios we consider three simulation areas. These are, in meters, 1000×1000, 1500×1500, and 2000×2000. For each situation, we simulate a network with 50 mobile nodes and let the simulation run for 5000 seconds. In the Disney World scenario, we choose node speeds that are realistic for walking, specifically, the minimum speed and the maximum speed are chosen to be 0.5 m/s and 2.0 m/s respectively. For the NYC scenario, we assume that people may take taxis or the subway and thus we allow the speed to vary from 15.0 m/s to 20.0 m/s. We now analyze three important characteristics of the network topologies generated by our mobility model.

A. Node Distribution

Figure 5 plots the locations of the nodes at times, in seconds, 1, 2500, and 5000. We only plot this for the Disney World scenario with a 1500×1500 simulation area as, in the other situations, the plots look very similar. Likewise, we looked at the plots for many other times and they also look very similar. We see that the nodes are evenly distributed in the simulation area at all times, and, in particular, there is no clustering of nodes in any one part of the simulation area. Since we start with a uniform distribution, it is not surprising that this holds at time 1, what is important is that this continues to hold for all of the other times. For a comparison, in the RWP model it is well-known that after a while the nodes start to cluster near the center.

B. Average Number of Neighbors

We say that two nodes are neighbors if they are close enough to each other that their wireless devices can communicate. For the purposes of our simulations we assume that the maximum node transmission range is 250 meters. Thus two nodes are neighbors if they are within 250 meters of each other. The average number of neighbors per node throughout the simulation is shown in Figure 6. For each simulation region we can see that the average number of neighbors is, essentially, constant. As we might expect, the number of neighbors decreases when the simulation area increases. This is because the node density decreases when the simulation area increases. The results further indicate that the average number of neighbors is similar between the NYC and the Disney World scenarios. This is most likely caused by the fact that, as we saw in Section VI-A, in both scenarios the nodes are evenly spread throughout the simulation area.

TABLE II. AVERAGE LINK DURATION (IN SECONDS).

Dimension	Disney	NYC
1000	295.32	38.01
1500	277.10	37.19
2000	259.08	36.28

C. Average Link Duration

When two nodes are neighbors we say that they have a link. If two nodes are neighbors, then they stop being neighbors, and then they become neighbors again, we consider this to be two different links. The amount of time that a link lasts is called the link duration. The average link duration in the system is an important factor that affects the connectivity of the network topology, and, as a result, it has a major impact on the performance of network protocols. In Table II we give the average link duration in each of the situations simulated. We see that the size of the simulation area does not affect the average link duration very much. On the other hand there is a huge difference in the average link duration for the Disney World and NYC scenarios. This is likely due to the differences in the speeds at which the nodes travel in these situations.

VII. CONCLUSIONS

In this paper we proposed a new mobility model called the Smoothly Truncated Lévy Walk (STLW). We presented a justification for its use from both a theoretical and an empirical perspective. We saw that it captures the characteristics of real-word traces very well, and, in particular, that modeling changes in direction is much more realistic than the standard approach of only modeling the direction of motion. Further, to better understand the topological properties of STLW we performed a series of simulations. We found that the node distribution and the number of neighbors do not vary much over time.

There are a number of ways to extend this model, which we leave for future work. One is to take into account some environmental information when available. For instance a simple way to model obstacles is as follows. If we know that a region in the simulation area cannot be crossed (perhaps it corresponds to a wall or a body of water) then we can treat this the same way that we treat a boundary, even though now it is within the simulation area. A further extension is to get a more realistic model for the speed. From the results in [19] it is clear that there is a relationship between speed and distance. The exact nature of this relationship is complicated and requires further study. An approach to modeling it is presented in [3]. It would be interesting to modify that approach to the present situation.

REFERENCES

[1] C. Bettstetter and C. Wagner. The spatial node distribution of the random waypoint mobility model. In *Mobile Ad-Hoc Netzwerke, deutscher Workshop über Mobile Ad-Hoc Netzwerke WMAN*, pages 41–58, 2002.

[2] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications & Mobile Computing (WCMC)*, 2:483–502, 2002.

[3] C.-H. Chen, H.-T. Wu, and K.-W. Ke. A novel mobility model of diverse average speed and uniform node distribution in mobile ad hoc networks. In *TENCON, IEEE Region 10 Conference*, 2006.

[4] P. K. Dunn. *tweedie: Tweedie exponential family models*, 2013. R package version 2.1.7.

[5] W. Feller. *An Introduction to Probability Theory and Its Applications Volume II*. John Wiley & Sons, Inc., New York, 2nd edition, 1971.

[6] M. C. González, C. A. Hidalgo, and A. L. Barabási. Understanding individual human mobility patterns. *Nature*, 453(7169):779–782, 2008.

[7] M. Grabchak. On a new class of tempered stable distributions: Moments and regular variation. *Journal of Applied Probability*, 49(4):1015–1035, 2012.

[8] M. Grabchak and S. Molchanov. Limit theorems and phase transitions for two models of summation of iid random variables depending on parameters. *Doklady Mathematics*, 88(1):431–434, 2013.

[9] M. Grabchak and S. Molchanov. Limit theorems and phase transitions for two models of summation of iid random variables with a parameter. To appear in *Theory of Probability and Its Applications*, 2014.

[10] M. Grabchak and G. Samorodnitsky. Do financial returns have finite or infinite variance? A paradox and an explanation. *Quantitative Finance*, 10(8):883–893, 2010.

[11] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy. Modeling spatial and temporal dependencies of user mobility in wireless mobile networks. *IEEE/ACM Transactions on Networking*, 17(5):1564–1577, 2009.

[12] A. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri. Towards realistic mobility models for mobile ad hoc networks. In *Proceedings of MOBICOM*, 2003.

[13] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In T. Imielinski and H. F. Korth, editors, *Mobile Computing*, pages 153–181. Springer, 1996.

[14] M. Kim, D. Kotz, and S. Kim. Extracting a mobility model from real user traces. In *Proceedings of INFOCOM*, pages 1–13, 2006.

[15] D. Kotz, T. Henderson, and I. Abyzov. CRAWDAD data set dartmouth/campus (v. 2004-12-18). Downloaded from http://crawdad.org/dartmouth/campus, 2004.

[16] U. Küchler and S. Tappe. Tempered stable distributions and processes. *Stochastic Processes and their Applications*, 123(12):4256–4293, 2013.

[17] M. Musolesi and C. Mascolo. Designing mobility models based on social network theory. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(3):59–70, 2007.

[18] D. A. Raichlen, B. M. Wood, A. D. Gordon, A. Z. P. Mabulla, F. W. Marlowe, and H. Pontzer. Evidence of Lévy walk foraging patterns in human hunter–gatherers. *Proceedings of the National Academy of Sciences of the Unites States of America*, 111(2):728–733, 2014.

[19] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong. On the levy-walk nature of human mobility: Do humans walk like monkeys? *IEEE/ACM Transaction on Networking*, 19(3):630–643, 2011.

[20] E. Royer, P. Melliar-Smith, and L. Moser. An analysis of the optimum node density for ad hoc mobile networks. In *IEEE International Conference on Communications*, pages 857–861, 2001.

[21] M. C. K. Tweedie. An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, pages 579–604, 1984.