Human Mobility Patterns and Their Impact on Routing in Human-Driven Mobile Networks

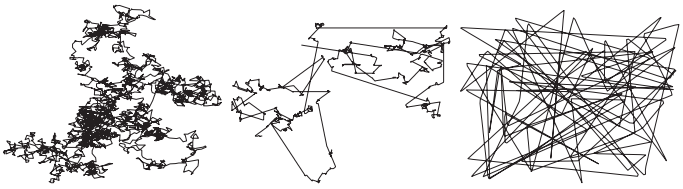
Injong Rhee (NCSU) Minsu Shin (NCSU) Seongik Hong (NCSU) Kyunghan Lee (KAIST) Song Chong (KAIST)

***Abstract*—We conduct a statistical study of human mobility using over 1000 hours of GPS traces of human walks involving 44 participants in five different locations, two university campuses, state fair, New York City, and Disney World. Our data reveals statistical features similar to those in what physicists have long called Levy random walks (or Levy walks). These features include heavy-tail distributions of flight lengths and super-diffusive nature of mobility. We find that these tendencies are likely caused by human intentions in deciding travel destinations (and distance and sojourn time thereof), but not by geographical constraints such as roads, buildings, boundaries, etc, and that geographical constraints, instead, cause truncations of flight lengths and discontinuity in the statistical tendency of mobility statistics. Based on these findings, we construct a simple Levy walk mobility model that emulates human walk patterns expected in outdoor mobile networks carried by humans. Based on these findings, we construct a simple Levy walk mobility model that emulates human walk patterns expected in outdoor mobile network environments. We demonstrate that the Levy walk model can be used to recreate the statistical patterns commonly observed in previous mobility studies such as the power-law distributions of human inter-contact times and that the simulation performance of mobile network routing protocols under the Levy walk model exhibits distinctive performance features unexplored under existing mobility models.**

I. Introduction

Biologists [19], [4], [14] have found that the mobility patterns of foraging animals such as spider monkey, albatrosses (seabirds) and jackals can be commonly described in what physicists have long called *Levy Walks*. The term Levy walks was first coined by Schlesinger et al. [17] to explain atypical particle diffusion not governed by Brownian motion (BM). BM characterizes the diffusion of tiny particles with a mean free path (or flight) and a mean pause time between flights. A *flight* is defined to be a longest straight line trip from one location to another that a particle makes without a directional change or pause. Einstein [8] first showed that the probability that such a particle is at a distance r from the initial position after time t has a Gaussian distribution and thus is proportional to $\sqrt{t}$, i.e., the width or standard deviation of a Gaussian distribution. The mean squared displacement (MSD), which is defined to be the variance of the probability distribution, is proportional to t. It is a manifestation of the central limit theorem (CLT) as the sum of flight lengths follows a Gaussian distribution. However, when flight lengths do not have a characteristic scale - in other words, their second moment is not finite, the particles are making Levy walks and may undergo atypical diffusion.

This work is supported in part by NSF CNS 0626850 and the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-352-D00137).



(a) (b) (c)

Fig. 1. Sample trajectories of (a) BM, (b) Levy walk and (c) RWP

This implies that the MSD of particles making Levy walks is proportional to t^γ where $\gamma > 1$ – thus CLT does not hold. Intuitively, Levy walks consist of many short flights and exceptionally long flights that eliminate the effect of such short flights in average flight lengths. The distribution of flights l in Levy walks is typically represented by an inverse power-law distribution: $p(l) \sim \frac{1}{l^{1+\alpha}}, 0 < \alpha < 2$ (note that BM has $\alpha \geq 2$). Sample trajectories of an object undergoing BM, Levy walks and RWP (random way point) are presented in Fig. 1 in which differences in the patterns are visually evident.

We study the statistical patterns of human walks observed within a radius of tens of kilometers. We use mobility track logs obtained from 44 participants carrying GPS receivers from September 2006 to January 2007. The sample settings where traces are obtained are two university campuses (one in Asia and one in the US), one metropolitan area (New York city), one State fair and one theme park (Disney World). The participants walk most of times in these locations and may also occasionally travel by bus, trolley, cars, or subway trains. These settings are selected because they are conducive to collecting GPS readings. Although the number of participants is relatively small in our study, the total duration of tracks taken over the five different sites are over 1000 hours, which adds to the statistical significance of our findings.

From the data analysis of our traces, we find the followings:

- The mobility patterns of the participants in these outdoor settings have features similar to those defining Levy walks; their flight distributions and pause time distributions closely match truncated power-law distributions. Their MSD also shows significant influence of these mobility patterns.
- There exist some deviations from pure Levy walks occurring due to various factors specific to human mobility including geographical constraints such as roads, buildings, obstacles and traffic. These deviations are manifested in our traces in the form of flight truncations which may make the flight distribution appear like heavy-tailed or even short-tailed at times.

Site (# of participants)	# of traces	Duration (hour)			Radius (km)		
		min	avg	max	min	avg	max
Campus I (20)	35	1.71	10.19	21.69	0.46	1.82	5.84
Campus II (4)	46	4.21	10.62	22.37	0.43	1.26	4.16
NYC (8)	30	1.23	9.34	22.66	0.37	4.18	6.98
DW (4)	15	4.43	8.68	13.20	0.39	1.67	4.43
SF (8)	8	1.81	2.57	3.12	0.22	0.28	0.34

TABLE I
STATISTICS OF COLLECTED MOBILITY TRACES FROM FIVE SITES.

To the best of our knowledge, this is the first work that studies the Levy walk nature of human walk mobility through real walk trace data, and none of the existing mobility models used for mobile network simulations captures the Levy walk characteristics of human walk mobility. Based on the statistical patterns obtained from the traces, we construct a simple Levy-walk model for use in mobile network simulations and show that the model can be used to create the power law inter-contact time distributions of human walks observed in [7]. We apply the Levy walk models to mobile network simulation and study the performance impact of Levy walks on routing performance in human-driven mobile networks including DTNs (delay-tolerant networks) and MANETs (mobile ad hoc networks). Our study reveals that compared to RWP, Levy walks induces much better MANET performance and much worse DTN performance. Given that many existing studies use RWP as the mobility model for the simulation of human-driven mobile networks, this result indicates that they have been greatly under-estimating MANET performance while greatly over-estimating DTN performance.

In this paper, we discuss these problems in more details and present some preliminary results. Our trace data and a more complete technical report will be available on-line.

II. MEASUREMENT METHODOLOGY

A. Data collection

Five sites are chosen for collecting human mobility traces. These are two university campuses (Campus I and Campus II), New York City, Disney World (Orlando), and North Carolina state fair. Garmin GPS 60CSx handheld receivers are used for data collection which are WAAS (Wide Area Augmentation System) capable with a position accuracy of better than three meters 95 percent of the time, in North America [1]. The GPS receivers take reading of their current positions at every 10 seconds and record them into a daily track log. The summary of daily traces is shown in Table I. The radius of each trace is a half of the maximum distance that a participant travels during a day.

All participants in the five scenarios are different individuals. We perform our analysis separately for each scenario and we do not aggregate traces from different scenario in our analysis. Nonetheless, as can be seen in the ensuing analysis, many statistical similarities are found among traces from different scenarios.

B. Trace analysis

From the traces, we extract the following data: flight length, pause time, direction, and velocity. To get these data from the traces, we map the traces into a two dimensional area (note that the GPS receivers produce three-dimensional positions), and to account for GPS errors, we clean the data as follows. We recompute a position at every 30 seconds by averaging three samples over that 30 second period (note GPS samples are taken at every 10 seconds). All the position information discussed below is based on the 30-second average positions.

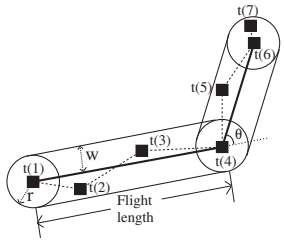


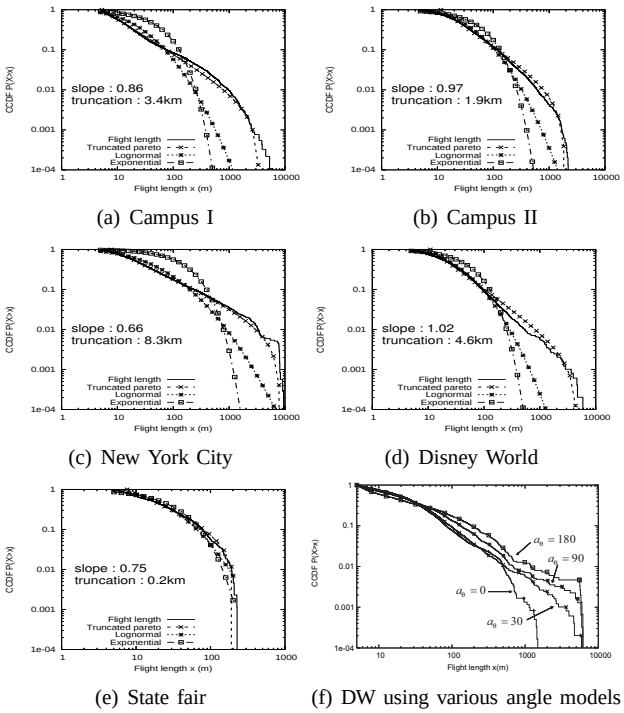
Fig. 2. The rectangular model used to extract flight information from traces.

As participants may move outside a line of sight from satellites or run out of battery, daily traces may contain discontinuities in time. For instance, if a participant disappears at time t (in seconds) at a position p from a trace and reappears at time $t + \Delta t$ at another position p', we use a similar method used in [12] to remove the discontinuity. If the next position recorded after the discontinuity is within a radius of 20 meters and the time to the next position is within a day boundary, then we assume that the participant walks to the next position from position p at a walking speed of 1 m/s from time $t + \Delta t - k$ (k is the distance between p and p' in meters) just before he shows up again at position p' in the trace and the remaining time ($\Delta t - k$) is recorded as a pause at the location where he disappeared. Otherwise, it is assumed that the trace has ended at time t and a new trace starts at time $t + \Delta t$.

We consider that a participant has a pause if the distance that he has moved during a 30 second period is less than r meters. It is not straightforward to extract flight information from a trace because people hardly move in a straight line. Combined with GPS errors, this human "errors" make it difficult to analyze flight data. To reduce noise due to these factors, we use three different methods, namely *rectangular*, *angle* and *pause-based* models. In the rectangular model, given two sampled positions x_s and x_e taken at times t and $t + \Delta t$ ($\Delta t > 0$) in the trace, we define the straight line between x_s and x_e to be a flight if and only if the following conditions are met: (a) the distance between any two consecutively sampled positions between x_s and x_e is larger than r meters (i.e., no pause during a flight), (b) when we draw a straight line from x_s to x_e, the sampled positions between these two end points are at a distance less than w meters from the line (the distance between the line and a position is the length of a perpendicular line from that position to the line) and (c) for the next sampled position x'_e after x_e, positions and the straight line between x_s and x'_e does not satisfy conditions (a) and (b). An example of the rectangular model is shown in Fig. 2. In that figure, the straight line movement between positions sampled at times $t(1)$ and $t(4)$ is regarded as one single flight between the two positions because all the sampled positions between them are inside of the rectangle formed by the two end points. In this example, the flight time is 90 seconds because each sample is taken at every 30 seconds. Both r and w are model parameters.

The angle model allows more flexibility in defining flights. In the rectangular model, a trip can be broken into small flights even though consecutive flights have similar directions. This implies even a small curvature on the road may cause multiple short flights. To remedy this, the angle model merges multiple successive flights acquired from the rectangular model into a single long flight if the following two conditions are satisfied: (a) no pause occurs between consecutive flights and (b) the relative angle (θ as shown in Fig. 2) between any two consecutive flights is less than a_θ degree. A merged flight is considered to be a straight line from the starting position of the first flight to the ending position of the last flight and its flight length is the length of that line. a_θ is a model parameter.

The pause-based model can be viewed as an extreme case of the angle model. The pause-based model merges all the successive flights from the rectangular model into a single flight if there is no pause between the flights. A merged flight is defined in the same way as in the angle model. This model produces significantly different trajectories from the actual GPS trajectories, due to the abstraction. However, it represents more faithfully human intentions to travel from one position to another without much deviation caused by geographical features such as roads, buildings and traffic.



(a) Campus I (b) Campus II

(c) New York City (d) Disney World

(e) State fair (f) DW using various angle models

Fig. 3. The CCDF of flight lengths ($a_\theta = 30$). Various known distributions are fitted using maximum likelihood estimation.

III. Human mobility

A power-law distribution of flight lengths is a hallmark of Levy walks. In this section, we study the distributions of flight lengths from our traces. In generating its distribution for each scenario, flight length samples from all the traces of the same site, regardless of their participants, are aggregated together and used in the same distribution. This "aggregation" is reasonable because every trace obtained from the same site is subject to the same or similar geographical constraints (i.e., roads, obstacles, traffic, and buildings). The same technique is used in other studies of Levy walks (e.g., [14]).

Fig. 3 shows the CCDF (complementary cumulative density function) of flight lengths from each scenario. CCDF is known to show the tail patterns of a distribution better than log-log binned PDF plots. The PDF plots can also be found in our full report. We apply Maximum Likelihood Estimation (MLE) to fit three known distributions, exponential, log-normal, and truncated Pareto distributions [3] to the CCDF. The MLE of the truncated Pareto is performed over the x-axis range between 50 meter and the 99.9% quantile of each distribution to isolate only the tail behavior. We observe that truncated Pareto has the best fit among the three distributions in all cases with truncation points over three-orders of magnitudes (i.e., 1000 meters), which is a rule of thumb for power-law distributions. To reduce the sensitivity on particular settings of our flight models, we vary the values of r, w and a_θ from 2.5 meters to 10 meters and from 15 degrees to 90 degrees, respectively. We performed line fitting on the tails of the resulting CCDF over several ranges.

Flight truncations are natural consequences of geographical constraints including boundaries and physical obstructions, and observation artifacts (e.g., we do not consider those flights that leave the area boundary). All the distributions in Fig. 3 suffer from truncations of flights longer than a few kilometers whose effects are shown as sharp drops in the frequency of very long flights. This effect show up evidently with State fair traces shown in Fig.3(e) where even short-tail distributions fit well. The State fair traces are obtained from a highly confined area of less than 350 meter radius (it is smallest among the five sites). Thus, it is subject to more truncations.

The sharp drops at the tails give rise to a possibility that the flight distributions have heavy-tails but not power tails since truncated power law distributions can be also fitted with non power-law heavy-tail distributions such as Weibull [9]. (This truncation problem also appears in earlier studies of animal mobility, e.g., [14].) Our data is inconclusive in disproving this. However, there are some hints that this may not be the case. Fig. 3(f) shows the CCDF of flights as we increase the flight angle in the flight model. We find that as the angle increases, the distribution becomes flatter with a heavier tail. Under the pause-based model (i.e., $a_\theta = 180$), it shows the heaviest tail. This phenomenon reveals an important feature in human mobility patterns: if we accept that humans tend to pause for a non-zero period of time when they get to a destination, the heavier-tail distribution of flights for the pause-based model implies that it is human intention causing the heavy-tail tendency, not the geographical constraints that force humans to make short flights with no pause (otherwise, the pause-based model should show almost the same tail distribution). This also implies the scale-free tendency of the flight distribution: as we increase the scale by removing constraints and boundaries or increasing the observation area size, we are expected to see longer flights. It does not make sense that human intention to move to a destination is bounded by some invisible boundaries as in Weibull (even though there

exists no physical bounds, e.g., building and campus). The power-law tendency of human mobility over a larger scale [6] also provides hints for this scale-freedom and self-similarity.

From the perspective of network simulations, power law distributions are easy to scale because simulation setups including geographical constraints may always change; for instance, simulation can run in a small area as well as a large area. It would be impossible to pick a different distribution for different setups. Using power-law distributions while inducing truncations as the natural consequence of adaptations to a given set of geographical constraints offers a much more convenient way of mobility simulation. Our data implies that although geographical constraints may vary in different scenarios, this scale-free tendency is invariant. For network simulations involving human-assisted mobile networks, while human navigation around obstacles and road shapes is relatively easy to program, the heavy-tail tendency of human intentions must be inherent in the mobility model to accurately depict human walk patterns.

IV. Levy-walk Mobility Model

In this section, we discuss a simple Levy-walk mobility model for simulating human carried mobile networks that generates synthetic mobility tracks reflecting the statistical patterns of human mobility that we find in our study. We use the same random walk model discussed in the section 2. A step is represented by four variables, flight length (l), direction (θ), flight time (Δt_f), and pause time (Δt_p). Our model picks flight lengths and pause times randomly from their PDFs $p(l)$ and $\psi(\Delta t_p)$ which are Levy distributions with coefficients α and β, respectively. The following defines a Levy distribution with a scale factor c and exponent α in terms of a fourier transformation,

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx - |ct|^\alpha} dt \qquad (1)$$

For $\alpha = 1$, it reduces to a Cauchy distribution and for $\alpha = 2$, a Gaussian with $\sigma = \sqrt{2}c$. Asymptotically, for $\alpha < 2$, $f_X(x)$ can be approximately by $\frac{1}{|x|^{1+\alpha}}$. We allow c, α and β to be simulation parameters.

Based on the above model, we generate synthetic Levy-walk mobility tracks with truncation factors τ_l and τ_p for flight lengths and pause times respectively in a confined area as follows. First, the initial location of a walker is picked randomly from a uniform distribution in the area. At every step, an instance of tuple $(l, \theta, \Delta t_f, \Delta t_p)$ is generated randomly from their corresponding distributions. If l and Δt_p are negative or $l > \tau_l$ or $\Delta t_p > \tau_p$, then we discard the step and regenerate another step. We repeat this process after the step time $\Delta t_f + \Delta t_p$. Until the end of the simulation, we generate the tuples repeatedly.

V. Routing Performance

In this section, we apply the mobility model developed in Section IV to the simulation of DTN and MANETs and measure routing performance in these networks and compare resulting routing patterns with those generated from existing models such as RWP and BM.

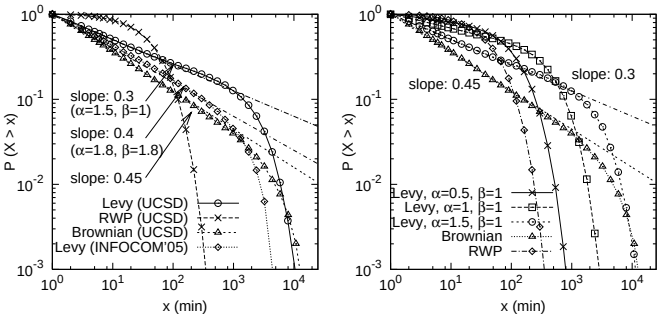
A. Routing in Delay Tolerant Networks

In delay tolerant networks (DTN), mobile nodes may establish on and off connectivity with their neighbors and the rest of the network. Therefore, store-and-forward is the main paradigm of routing in such networks where communication transpires only when two devices are in a radio range. We call the time period that two nodes are in a radio range the *contact time* of the two nodes. One of the most widely studied routing algorithms in DTN is *two-hop relay routing* [18] where a source node sends a message (or a sequence of data packets) to the first node it contacts and then that first node acts as a relay and delivers the message when it contacts the destination node of the message. Here the period between the time that the message has originated and the time that the message is delivered to the relay node is called *first contact time* (FCT) and the period after that to the time the message is delivered to the destination is called *remaining inter-contact time* (RICT). In a dense network, FCT is typically negligible and RICT dominates the message delay. One way to characterize RICT is to measure the *inter-contact time* (ICT), the time period between two successive contact times of the same two nodes. Since it is difficult to measure RICT from real mobility traces, ICT has been used to characterize RICT [7].

It is known that the ICT of human mobility exhibits a strong power-law tendency [7]. The result is interesting because [16] showed by simulation that RWP produces exponentially decaying ICT, implying human mobility cannot be modeled by RWP. What's not obvious is the type of mobility patterns that gives rise to the power-law tendency of ICT distributions. In this section, we explore this problem using the mobility model from Section IV.

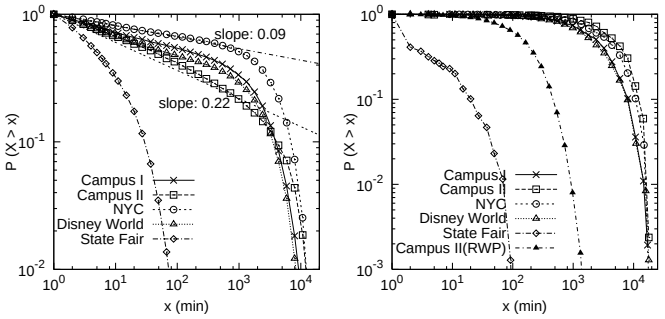
The earlier measurement studies on ICT (e.g., [7]) report power-law distributions of ICT with human mobility with slopes in the range of [0.3,0.4]. By varying the parameters of α and β of our mobility model, we are able to generate ICT distributions with the similar characteristics as in [7] by MATLAB simulation. [7] reports power-law slopes of 0.3 from the INFOCOM trace [10] and 0.4 from the UCSD trace [13]. Fig. 4(a) shows the result. In the UCSD simulation, we fix the simulation area to 3.5 km by 3.5 km, τ_l to 3 km and τ_p to 28 hours. These values are chosen based on the data from [13]. The transmission range of each node is set to 250 meter radius (which is typical for IEEE 802.11b). For the INFOCOM simulation, we set the area to 1.5 km by 1.5 km, τ_l to 200 m, τ_p to 1 hour and the transmission range of each node to 100 m – the maximum transmission range for the Bluetooth devices used for taking the original traces. 40 nodes are simulated in both scenarios for 300 hours. For all the simulations, we assume infinite buffer and that message transfers occur instantaneously. These assumptions are used to isolate the effect of mobility patterns on the performance of DTN routing.

We also simulate RWP and BM in the same setup as the UCSD environment to compare the results. BM's ICT distribution shows 0.45 power-law slope while RWP's shows an exponential decay. Although there could be other types of mobility patterns that could generate the same ICT distri-



(a) (b)

Fig. 4. The ICT distributions of mobility models. Levy walks recreate the ICT distributions seen in the INFOCOM and UCSD traces.

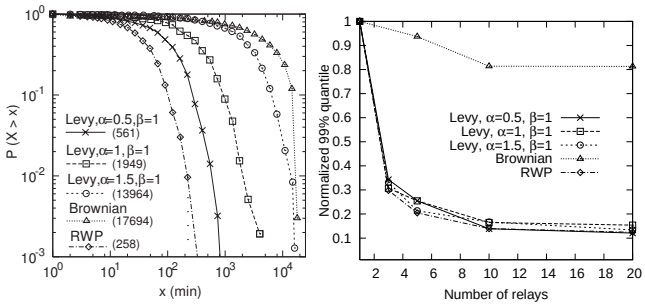


(a) CCDF of ICT (b) CCDF of routing delays

Fig. 5. ICT and DTN routing delays under various models including those constructed based on the statistics from our traces.

butions as INFOCOM's and UCSD's, this result allows us to conjecture that the actual mobility that generates these characteristics in these settings might have been Levy walks. The more diffusive the mobility is, the shorter tail its ICT distribution becomes. To confirm this pattern, we run Levy walks with different values of α while fixing β to one. Fig. 4(b) shows that as α gets smaller, the tail distribution of ICT becomes shorter.

ICT directly impacts routing delays in DTN. To see their relation, we measure the ICT and routing delays from simulation setups that mimic the environments of our five scenarios. In this simulation, we set α and β to the average values extracted from the angle model (with $a_\theta = 30$ and $r = 5$), set the scale factors for flight lengths and pause times to 10 and 1, respectively, and match the simulation areas to the same as those of the five sites and the flight length and pause time truncation points are set to those measured from the traces. The routing delays of RWP which uses the same environment as Campus II are also measured for comparison. In all scenarios, we simulate 40 nodes. The resulting routing delay distributions along with their ICT's are shown in Fig. 5. The figure shows that all the ICT distributions from our models, except that from State fair, follow strong power-law for the duration up to several hours. The truncations in the ICT distributions occur because the simulation is cut off around 300 hours and we consider only those contacts made within the simulation time. The ICT distribution from State fair exhibits exponential decay. This is because the area of State fair (a radius of 340 m) is much smaller than the others, and given the transmission range of 250 m, nodes can make contacts with each other without traveling a long distance. The routing delays of the corresponding Levy walk models tend to have high delays because, among many factors influencing the delays, the simulation area of our models is particularly large: on average, our models have at least four times larger an area than the UCSD simulation area. However, the routing delays of RWP still show a short tail distribution. To see the effect of flight length distributions on routing delays more clearly, we measure routing delays in the simulation runs used for Fig. 4 (b). Figure 6 (a) shows the result from which the following can be observed. BM tends to have much larger delays than any other models while RWP, as expected, shows the smallest delays because its probability of long flights is highest. The

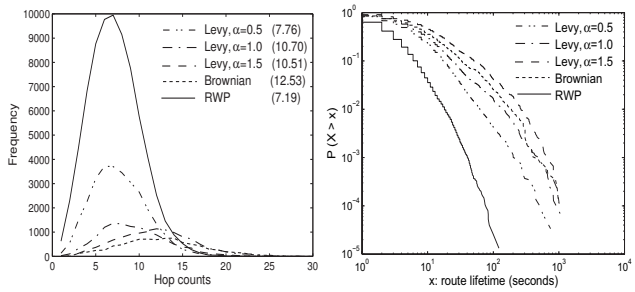


(a) One relay (b) Multiple relays

Fig. 6. The DTN delay distributions of various mobility models and normalized 99% quantile delay with multiple relays. The numbers in the parenthesis represent the actual delays in minutes at the 99% quantile of the distributions.

Levy walk models show their patterns in between the two extremes: as we increase α, their delays get closer to BM's and as we reduce α, they get closer to RWP.

The heavy tail distribution of routing delays may intuitively imply that many nodes experience similar long routing delays and that use of more relays (or copies of messages) may not necessarily improve the performance drastically. In a generalized relaying algorithm, the source distributes the message to the first m relays that it contacts. The routing delay is the time till any copy of the message is delivered to the destination. Fig. 6 (a) shows the DTN routing delays of various models when one relay is used, and Fig. 6 (b) shows the 99% quantile delays of the same models normalized by their corresponding one-relay delays as we add more relays. As expected, BM hardly achieves this goal; the delay does not improve so much as the number of relays increases, since every relay takes long time to meet the destination. However, we are surprised to find that all our Levy walk models including the one with $\alpha = 1.5$ which shows fairly similar delay patterns as BM for one relay case, show almost the same improvement ratio as RWP as we add more relays. This implies that while in RWP, most nodes travel long distances frequently, in Levy walks, although not all nodes make such long trips, there exist with high probability some nodes within the mobility range of the source nodes that make such long trips. This contributes to the great reduction of the delays even with a small number of relays.



(a) (b)

Fig. 7. (a) The hop count distributions of the shortest path between two randomly selected nodes undergoing various mobility patterns. The numbers inside the parenthesis represent the average hop counts. (b) the CCDF of their corresponding path durations.

B. Routing in MANETs

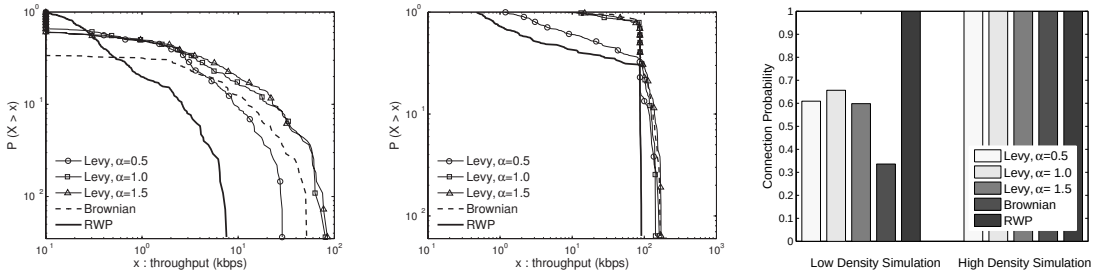
In this section, we examine the impact of Levy walks on the performance of MANET routings. There exist many MANET protocols in the literature. It is impractical to evaluate all the protocols, but instead, we first focus on the features of mobility that affects the performance of MANET routing such as hop counts and path durations. These features strongly influence the routing performance of MANETs. For instance, [15] shows that data throughput is proportional to path durations within the limit of link capacity in the network.

Fig. 7 shows the hop count distributions of the shortest path between two randomly picked nodes in the simulation of various mobility models, and the CCDF of their corresponding path durations. We use the same simulation setup as discussed in Section IV. The radio range of each mobile is set to 250 meter. We run the simulation for 3000 seconds. 400 pairs of nodes are selected and the hop count of each pair is measured and sampled once at each time they establish a new path. Most of RWP hop counts are less than 15 hops and their distribution is peaked around 7 hops. This occurs because RWP nodes tend to cluster around the center of the simulation area [5]. The hop count samples of RWP are also much larger than those of the other models because as we can see in Fig. 7(b), RWP tends to maintain much shorter routing paths than the other models because of its high mobility. On the other hand, Levy walks tend to have longer paths than RWP. Because of the less diffusive nature of Levy walks, Levy walk nodes tend to stay longer in one location than RWP. Therefore, nodes are more spread out in the simulation area than RWP. Since the path durations of Levy walks are longer than those of RWP, the numbers of hop count samples of these models are much smaller. BM shows an extreme case of inactivity as its average hop count is longest. However, its path duration is the second longest to the Levy walks model with $\alpha = 1.5$. This is because most paths with long durations are from short paths, and BM nodes are more spread out and tend to have less chance of short paths as we can see from Fig. 7(a). In addition, because BM has $\beta = 2$, it has more occurrences of short pause times than the Levy walk models. These factors collectively contribute to reducing the path durations of BM below that of the Levy walk model with $\alpha = 1.5$ although BM is slightly less diffusive than the Levy walk model.

To see the effects of the above-discussed factors on routing performance, we simulate DSR [11], a source-based MANET routing, in the same simulation setup as the above using GloMoSim [2]. In this simulation, we measure the data throughput of FTP connections over 300 node pairs randomly selected. The link bandwidth in these simulations is set to 2 Mbps. Figs. 8 (a) and (b) show the CCDF of throughput measured in low and high node density network environments for various mobility models. For the high density environment, we use 100 nodes in 1 km by 1 km area with τ_l = 500 m and for the low density environment, 2 km by 2 km area with τ_l = 1 km. We use the same values for the other parameters as in the simulation run for Fig. 7.

In general, both hop counts and path durations have significant impact on routing throughput. Typically, the influence of hop counts itself on data throughput gets less emphatic as hop counts increase because each simulation run contains one connection so there is only self-interference, and self-interference is limited only within a few hops. However, it is clear that as the number of hops of a path increases, its path duration is likely to reduce. Path durations seem to be a significant determinant of data throughput in our simulation. This can be seen from the similarity of Figs. 7(b) and 8(a).

In the low density simulation, the node pairs with the best throughput around the tail of the throughput CCDF tend to have long path durations. In the simulation BM and Levy walks have an order of magnitude higher maximum throughput than RWP. However, around the top of the CCDF in the figure, BM and Levy walks show a significantly less number of node pairs. This is because the number of successful path connections is much less for BM and Levy walks. In Fig. 8(c), we plot the connection probability of node pairs, the probability that two randomly selected nodes successfully establish a routing path between them, in our simulation runs. The connection probabilities of BM and Levy walks are around 30% and 60%, respectively. This is because the difference in the diffusion rates of mobility has influenced their clustering behavior. As mentioned above, BM and Levy walk nodes tend to be more spread out, likely incurring more disconnected islands. On the other hand, while RWP nodes do not have any connectivity problem, their throughput tends to be much lower than that of BM and Levy walks. These factors collectively cause BM and Levy walks to have heavier tail throughput distributions while causing RWP to have a short tail. Thus, when examining network performance under realistic mobility models, we need to examine the entire distribution of performance instead of single numbers such as average or median values which are much less meaningful under power-law distributions of performance metrics of interest. Under the high density network simulation, all mobility models achieve 100% connection probability. Even in this environment, the data throughput under BM and Levy walks is much higher than that of RWP because of their longer path durations.



(a) (b) (c)

Fig. 8. (a) CCDF of FTP throughput in a low node density simulation (b) CCDF of FTP throughput in a high node density simulation (c) the probability of establishing a route between two randomly selected nodes under the low and high node density simulations

VI. Conclusion

Levy walks are commonly observed in biology and physics and they are also optimal ways for foraging animals to find randomly dispersed food sources in optimal ways. If we can leverage these lessons learned from nature for DTN routing problems, good strategies to evaluate the performance of routing and also to route messages as well are likely obtained.

References

[1] Garmin GPSMAP 60CSx User's manual. http://www.garmin.com/products/gpsmap60csx/.

[2] Glomosim from UCLA homepage. http://pcl.cs.ucla.edu/projects/glomosim/.

[3] I. B. Aban, M. M. Meerschaert, and A. K. Panorska. Parameter estimation for the truncated pareto distribution. *Journal of the American Statistical Assoc.*, 101(473):270–277, March 2006.

[4] R. P. D. Atkinson, C. J. Rhodes, D. W. Macdonald, and R. M. Anderson. Scale-free dynamics in the movement patterns of jackals. *OIKOS*, 98(1):134–140, 2002.

[5] C. Bettstetter, G. Resta, and P. Santi. The node distribution of the random waypoint mobility model for wireless ad hoc networks. *IEEE Trans. Mobile Computing*, 2(3):257–269, July 2003.

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

[7] 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 *Proc. of IEEE INFOCOM 2006*, Spain, April 2006.

[8] A. Einstein. On the motion, required by the molecular-kinetic theory of heat, of particles suspended in a fluid at rest. *Ann. Phys.*, 17:549–560, 1905.

[9] A. Feldmann and W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. In *Proc. of IEEE INFOCOM'97*, Kobe, Japan, April 1997.

[10] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. In *Proc. of ACM WDTN '05*, pages 244–251, Philadelphia, PA, August 2005.

[11] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In Imielinski and Korth, editors, *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.

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

[13] M. McNett and G. M. Voelker. Access and mobility of wireless pda users. *SIGMOBILE Mob. Comput. Commun. Rev.*, 9(2):40–55, 2005.

[14] G. Ramos-Fernandez, J. L. Morales, O. Miramontes, G. Cocho, H. Larralde, and B. Ayala-Orozco. Levy walk patterns in the foraging movements of spider monkeys (ateles geoffroyi). *Behavioural Ecology and Sociobiology*, 55:223–230, 2004.

[15] N. Sadagopan, F. Bai, B. Krishnamachari, and A. Helmy. Paths: analysis of path duration statistics and their impact on reactive manet routing protocols. In *Proc. of ACM MobiHoc '03*, pages 245–256, June 2003.

[16] G. Sharma and R. R. Mazumdar. Scaling laws for capacity and delay in wireless ad hoc networks with random mobility. In *Proc. of IEEE ICC 2004*, Paris, France, June 2004.

[17] M. F. Shlesinger, J. Klafter, and Y. M. Wong. Random walks with infinite spatial and temporal moments. *J. Stat. Phys.*, 27:499–512, 1982.

[18] K. F. Sushant Jain and R. Patra. Routing in a delay tolerant network. In *SIGCOMM*, SIGCOMM, 2004.

[19] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley. Levy flights search patterns of wandering albatrosses. *Nature*, 381:413–415, 1996.