Explaining the power-law distribution of human mobility through transportation modality decomposition

Kai Zhao and Sasu Tarkoma
Department of Computer Science, University of Helsinki, Helsinki, Finland

Mirco Musolesi
School of Computer Science, University of Birmingham, Birmingham, UK

Pan Hui
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Hong Kong

Weixiong Rao
School of Software Engineering, Tongji University, Shanghai, China

Human mobility has been empirically observed to exhibit Lévy flight characteristics and behaviour with power-law distributed jump size. The fundamental mechanisms behind this behaviour has not yet been fully explained. In this paper, we propose to explain the Lévy walk behaviour observed in human mobility patterns by decomposing them into different classes according to the different transportation modes, such as Walk/Run, Bike, Train/Subway or Car/Taxi/Bus. We show that human mobility can be modelled as a mixture of different transportation modes, and that these single movement patterns can be approximated by a lognormal distribution rather than a power-law distribution. Then, we demonstrate that the mixture of the decomposed lognormal flight distributions associated with each modality is a power-law distribution, providing an explanation to the emergence of Lévy Walk patterns that characterize human mobility patterns.
Understanding human mobility is crucial for epidemic control\textsuperscript{1,4} urban planning\textsuperscript{8,6}, traffic forecasting systems\textsuperscript{5,3} and, more recently, various mobile and network applications\textsuperscript{9,13}. Previous research has shown that trajectories in human mobility have statistically similar features as Lévy Walks by studying the traces of bank notes\textsuperscript{14}, cell phone users’ locations\textsuperscript{12} and GPS\textsuperscript{10,19}. According to this model, human movement contains many short flights and some long flights, and these flights follow a power-law distribution.

Intuitively, these long flights and short flights reflect different transportation modalities. Figure\textsuperscript{1} shows a person’s one-day trip with three transportation modalities in Beijing based on the Geolife dataset (Table\textsuperscript{1} 20\textsuperscript{22}). Starting from the bottom right corner of the figure, the person takes a taxi and then walks to the destination in the top left part. After two hours, the person takes the subway to another location (bottom left) and spends five hours there. Then the journey continues and the person takes a taxi back to the original location (bottom right). The short flights are associated with walking and the second short-distance taxi trip, whereas the long flights are associated with the subway and the initial taxi trip. Based on this simple example, we observe that the flight distribution of each transportation mode is different.

In this paper, we study human mobility with two large GPS datasets (approximately 10 and 20 million GPS samples), both containing transportation mode information such as Walk/Run, Bike, Train/Subway or Car/Taxi/Bus. The four transportation modes (Walk/Run, Bike, Train/Subway and Car/Taxi/Bus) cover the most frequently used human mobility cases. First, we simplify the trajectories obtained from the datasets using a rectangular model, from which we obtain the flight length\textsuperscript{16}. Here a flight is the longest straight-line trip from one point to another without change of direction\textsuperscript{16,19}. One trail from an origin to a destination may include several different flights (Fig.\textsuperscript{1}). Then, we determine the flight length distributions for different transportation modes. We fit the flight distribution of each transportation mode according to the Akaike information criteria\textsuperscript{23} in order to find the best fit distribution.

We show that human movement exhibiting different transportation modalities is better fitted with the lognormal distribution rather than the power-law distribution. Finally, we demonstrate that the mixture of these transportation mode distributions is a power-law distribution based on two new findings: first, there is a significant positive correlation between consecutive flights in the same transportation mode, and second, the elapsed time in each transportation mode is exponentially distributed.

The contribution of this paper is twofold. First, we extract the distribution function of displacement with different transportation modes. This is important for many applications\textsuperscript{7,11}. Second, we demonstrate that the mixture of different transportation can be approximated as a truncated Lévy Walk. This result is a step towards explaining the emergence of Lévy Walk patterns in human mobility.

## RESULTS

**Power-law fit for overall flight.** First, we fit the flight length distribution of the Geolife and Nokia MDC datasets regardless of transportation modes. We fit truncated power-law, lognormal, power-law and exponential distribution (see Supplementary Table S1). We find that the overall flight length (\(l\)) distributions fit a truncated power-law \(P(l) \propto l^\alpha e^{-\lambda l}\) with exponent \(\alpha\) as 1.57 in the Geolife dataset (\(\lambda = 0.00025\)) and 1.39 in the Nokia MDC dataset (\(\lambda = 0.00016\)) (Fig.\textsuperscript{2}), better than other alternatives such as power-law, lognormal or exponential. Figure\textsuperscript{2} illustrates the PDFs and their best fitted distributions according to Akaike weights. The best fitted distribution (truncated power-law) is represented as a solid line and the rest are dotted lines. We use logarithm bins to remove tail noises\textsuperscript{16,24}. Our result is consistent with previous research (\textsuperscript{14,19}), and the exponent \(\alpha\) is close to their results.

We show the Akaike weights for all fitted distributions in the Supplementary Table S2. The Akaike weight is a value between 0 and 1. The larger it is, the better the distribution is fitted\textsuperscript{23,24}. The Akaike weights of the power-law distributions regardless of transportation modes are 1.0000 in both datasets. The \(p\)-value is less than 0.01 in all our tests, which means that our results are very strong in terms of statistical significance.

**Lognormal fit for single transportation mode.** However, the distribution of flight lengths in each single transportation mode is not well fitted with the power-law distribution. Instead, they are better fitted with the lognormal distribution (see Supplementary Table S2). All the segments of each transportation flight length are best approximated by the lognormal distribution with different parameters. In Fig.\textsuperscript{3} and Supplementary Fig. S1, we represent the flight length distributions of Walk/Run, Bike, Subway/Train and Car/Taxi/Bus in the Geolife and the Nokia MDC dataset correspondingly. The best fitted distribution (lognormal) is represented as a solid line and the rest are dotted lines.

Table\textsuperscript{2} shows the fitted parameter for all the distributions (\(\alpha\) in the truncated power-law, \(\mu\) and \(\sigma\) in the lognormal). We can easily find that the \(\mu\) is increasing over these transportation modes (Walk/Run, Bike, Car/Taxi/Bus and Subway/Train), identifying an increasing average distance. Compared to Walk/Run, Bike or Car/Taxi/Bus, the flight distribution in Subway/Train mode is more right-skewed, which means that people usually travel to a more distant location by Subway/Train.
It must be noted that our findings for the Car/Taxi/Bus mode are different from these recent research results [25, 26], which also investigated the case of a single transportation mode, and found that the scaling of human mobility is exponential by examining taxi GPS datasets. The differences are mainly because few people tend to travel a long distance by taxi due to economic considerations. So the displacements in their results decay faster than those measured in our Car/Taxi/Bus mode cases.

**Mechanisms behind the power-law pattern.**

We characterize the mechanism of the power-law pattern with Lévy flights by mixing the lognormal distributions of the transportation modes. Previous research has shown that a mixture of lognormal distributions based on an exponential distribution is a power-law distribution [29, 32]. Based on their findings, we demonstrate that the reason that human movement follows the Lévy Walk pattern is due to the mixture of the transportation modes they take.

We demonstrate that the mixture of the lognormal distributions of different transportation modes (Walk/Run, Bike, Train/Subway or Car/Taxi/Bus) is a power-law distribution given two new findings: first, there is a strong correlation between consecutive flights in the same transportation mode (the change rate in the same transportation mode is small over time), and second, the elapsed time between different transportation modes is exponentially distributed.

**Lognormal in the same transportation mode.** Suppose the flight length \( l_t \) is the flight length at time \( t \). The flight length at next interval of time \( l_{t+1} \), given the change rate \( c_{t+1} \), is

\[
l_{t+1} = l_t + c_{t+1} l_t.
\]

(1)

It has been found that the change rate \( c_t \) in the same transportation mode is small over time [4, 22]. The change rate \( c_t \) reflects the correlation between two consecutive displacements in one trip. To obtain the pattern of correlation between consecutive displacements in each transportation mode, we plot the flight length point \( (l_t, l_{t+1}) \) from the GeoLife dataset (Fig. S2). Here \( l_t \) represents the \( t \)-th flight length and \( l_{t+1} \) represents the \( t + 1 \)-th flight length in a consecutive trajectory in one transportation mode [33]. Figure. S2 shows the density of flight lengths correlation in Car/Taxi/Bus, Walk/Run, Subway/Train and Bike correspondingly. \( (l_t, l_{t+1}) \) are posited near the diagonal line, which identifies a clear positive correlation. Similar results are also found in the Nokia MDC dataset (see Supplementary Fig. S2).

We use the Pearson correlation coefficient to quantify the strength of the correlation between two consecutive flights in one transportation mode [34]. The value of Pearson correlation coefficient \( r \) is shown in the Supplementary Table S3. The \( p \) value is less than 0.01 in all the cases, identifying very strong statistical significances. \( r \) is positive in each transportation mode and ranges from 0.3640 to 0.6445, which means that there is a significant positive correlation between consecutive flights in the same transportation mode, and the difference \( c_t \) in the same transportation mode between two time steps is small.

The difference \( c_t \) in the same transportation mode between two time steps is small due to a small difference \( l_{t+1} - l_t \) in consecutive flights. We sum all the contributions as follows:

\[
\sum_{t=0}^{T} c_t = \sum_{t=0}^{T} \frac{l_{t+1} - l_t}{l_t} \approx \int_{0}^{T} \frac{dl}{l} = \ln \frac{l_T}{l_0}. 
\]

(2)

(3)

By the Central Limit Theorem, the sum of \( c_t \) is normally distributed over the unit interval, with mean \( \mu \) and variance \( \sigma^2 \). For every time step \( t \), the logarithm of \( l \) is also normally distributed with a mean \( \mu \) and variance \( \sigma^2 \) [35]. This means that the distribution of the flight length of the same transportation mode is lognormal, its density is given by

\[
P_{\text{singlemode}}(l) = \frac{1}{l \sqrt{2\pi \sigma^2 l}} \exp\left(-\frac{\ln(l) - \mu l}{2\sigma^2 l}\right),
\]

(4)

which corresponds to our findings that in each single transportation mode the flight length is lognormal distributed.

**Exponential elapsed time.** We also find that the elapsed time \( t \) is weighted exponentially between the different transportation modes (see Supplementary Fig. S3). Similar results are also reported in [25, 36]. Here the elapsed time \( t \) is the time taken under different transportation modes. For example, the trajectory samples shown in Fig. 1 contains six trajectories with three different transportation modes, (taxi, walk, subway, walk, taxi, walk). Thus the elapsed time also consists of six samples \( (t_{\text{taxi1}}, t_{\text{walk1}}, t_{\text{subway1}}, t_{\text{walk2}}, t_{\text{taxi2}}, t_{\text{walk3}}) \). The exponentially weighted time interval is mainly due to a large portion of Walk/Run flight intervals. Walk/Run is usually a connecting mode between different transportation modes, and Walk/Run usually takes much shorter time than any other modes. Thus the elapsed time decays exponentially.
Mixture of the transportation modes. Given these lognormal distributions \( P_{\text{single mode}}(l) \) in each transportation mode and the exponential elapsed time \( t \) between different modes, we make use of mixtures of distributions. If the distribution of \( x \), \( p(x,t) \), depends on the parameter \( t \), \( t \) is also distributed according to its own distribution \( r(t) \). Then the distribution of \( x \), \( p(x) \) is given by \( p(x) = \int_0^\infty p(x,t)r(t)dt \).

So the mixture (overall flight length \( P_{\text{overall}}(l) \)) of these lognormal distributions in one transportation mode given an exponential elapsed time (with an exponent \( \lambda \)) between each transportation mode is

\[
P_{\text{overall}}(l) = \int_0^\infty \lambda exp(-\lambda t) \frac{1}{\sqrt{2\pi \sigma^2 t}} \exp\left[-\frac{(\ln(l) - \mu)^2}{2\sigma^2 t}\right]dt,
\]

which can be calculated to give

\[
P_{\text{overall}}(l) = Cl^{-\alpha'},
\]

where the power law exponent \( \alpha' \) is determined by \( \alpha' = 1 - \frac{\mu^2 + 2\lambda\sigma^2}{\sigma^2} \). If we substitute the parameters presented in Table 2 we will get the \( \alpha' = 1.55 \) in the Geolife dataset, which is close to the original parameter \( \alpha = 1.57 \), and \( \alpha' = 1.40 \) in the Nokia MDC dataset, which is close to the original parameter \( \alpha = 1.39 \). The result verifies that the mixture of these correlated lognormal distributed flights in one transportation mode given an exponential elapsed time between different modes is a truncated power-law distribution. The calculation to obtain \( \alpha' \) is given in Supplementary Note 1.

DISCUSSION

Previous research suggests that it might be the underlying road network that governs the Lévy flight human mobility, by exploring the human mobility and examining taxi traces in one city in Sweden 19. To verify their hypothesis, we use a road network dataset of Beijing containing 433,391 roads with 171,504 conjunctions and plot the road length distribution 16, 27, 28. As shown in Supplementary Fig. S4, the road length distribution is very different to our power-law fit in flights distribution regardless of transportation modes. The \( \alpha \) in road length distribution is 3.4, much larger than our previous findings \( \alpha = 1.57 \) in the Geolife and \( \alpha = 1.39 \) in the Nokia MDC. Thus the underlying street network cannot fully explain the Lévy flight in human mobility. This is mainly due to the fact that it does not consider many long flights caused by metro or train, and people do not always turn even if they arrive at a conjunction of a road. Thus the flight length tails in the human mobility should be much larger than those in the road networks.

METHODS

Data Sets. We use two large real-life GPS trajectory datasets in our work, the Geolife dataset 20 and the Nokia MDC dataset 37. The key information provided by these two datasets is summarized in Table 1. We extract the following information from the dataset: flight lengths and their corresponding transportation modes.

Geolife 20, 22 is a public dataset with 182 users’ GPS trajectory over five years (from April 2007 to August 2012) gathered mainly in Beijing, China. This dataset contains over 24 million GPS samples with a total distance of 1,292,951 kilometers and a total of 50,176 hours. It includes not only daily life routines such as going to work and back home in Beijing, but also some leisure and sports activities, such as sightseeing, and walking in other cities. The transportation mode information in this dataset is manually logged by the participants.

The Nokia MDC dataset 37 is a public dataset from Nokia Research Switzerland that aims to study smartphone user behaviour. The dataset contains extensive smartphone data of two hundred volunteers in the Lake Geneva region over one and a half years (from September 2009 to April 2011). This dataset contains 11 million data points and the corresponding transportation modes.

Obtaining Transportation Mode and The Corresponding Flight Length. We categorize human mobility into four different kinds of transportation modality: Walk/Run, Car/Bus/Taxi, Subway/Train and Bike. The four transportation modes cover the most frequently used human mobility cases. To the best of our knowledge, this article is the first work that examines the flight distribution with all kinds of transportation modes in both urban and intercity environments. In the Geolife dataset, users have labelled their trajectories with transportation modes, such as driving, taking a bus or a train, riding a bike and walking. There is a label file storing the transportation mode labels in each user’s folder, from which we can obtain the ground truth transportation mode each user is taking and the corresponding timestamps. Similar to the Geolife dataset, there is also a file storing the transportation mode with an activity ID in the Nokia MDC dataset. We treat the transportation mode information in these two datasets as the ground truth.
In order to obtain the flight distribution in each transportation mode, we need to extract the flights. We define a flight as the longest straight-line trip from one point to another without change of direction [16, 19]. One trail from an original to a destination may include several different flights (Fig. 1). In order to mitigate GPS errors, we recompute a position by averaging samples (latitude, longitude) every minute. Since people do not necessarily move in perfect straight lines, we need to allow some margin of error in defining the ‘straight’ line. We use a rectangular model to simplify the trajectory and obtain the flight length [16]. We map the flight length with transportation modes according to timestamp in the Geolife dataset and activity ID in the Nokia MDC dataset and obtain the final (transportation mode, flight length) patterns. We obtain 202,702 and 224,723 flights with transportation mode knowledge in the Geolife and Nokia MDC dataset, respectively.

**Identifying the Scale Range.** To fit a heavy tailed distribution such as a power-law distribution, we need to determine what portion of the data to fit ($x_{\text{min}}$) and the scaling parameter ($\alpha$). We use the methods from [29, 38] to determine $x_{\text{min}}$ and $\alpha$. We create a power-law fit starting from each value in the dataset. Then we select the one that results in the minimal Kolmogorov-Smirnov distance, between the data and the fit, as the optimal value of $x_{\text{min}}$. After that, the scaling parameter $\alpha$ in the power-law distribution is given by

$$\alpha = 1 + n \left( \sum_{i=1}^{n} \ln \frac{x_i}{x_{\text{min}}} \right)^{-1},$$

(7)

where $x_i$ are the observed values of $x_i > x_{\text{min}}$ and $n$ is number of samples.

**Akaike weights.** We use Akaike weights to choose the best fitted distribution. An Akaike weight is a normalized distribution selection criterion [23]. Its value is between 0 and 1. The larger the value is, the better the distribution is fitted.

Akaike’s information criterion (AIC) is used in combination with Maximum likelihood estimation (MLE). MLE finds an estimator of $\hat{\theta}$ that maximizes the likelihood function $L(\theta|\text{data})$ of one distribution. AIC is used to describe the best fitting one among all fitted distributions,

$$AIC = -2\log \left( L(\hat{\theta}|\text{data}) \right) + 2K.$$

(8)

Here $K$ is the number of estimable parameters in the approximating model.

After determining the AIC value of each fitted distribution, we normalize these values as follows. First of all, we extract the difference between different AIC values called $\Delta_i$,

$$\Delta_i = AIC_i - AIC_{\text{min}}.$$

(9)

Then Akaike weights $W_i$ are calculated as follows,

$$W_i = \frac{\exp(-\Delta_i/2)}{\sum_{r=1}^{R} \exp(-\Delta_r/2)}.$$

(10)
[36] Zhu, H. et al. Recognizing exponential inter-contact time in vanets. In *INFOCOM*, 101–105 (2010).

[37] Kiukkonen, N., J., B., Dousse, O., Gatica-Perez, D. & J., L. Towards rich mobile phone datasets: Lausanne data collection campaign. In *Proc. ACM Int. Conf. on Pervasive Services (ICPS), Berlin*. (2010).

[38] Clauset, A., Shalizi, C. & Newman, M. Power-law distributions in empirical data. *SIAM Review* **51**, 661–703 (2009).
Figure 1. Illustration of a synthetic trail (taxi, walk, subway, walk, taxi, walk) for one day trip and their corresponding flights. This figure shows that the flight distribution of each transportation mode (walk, taxi, subway) is very different.
Figure 2. Power-law fit for overall flight. (a-b) Power-law fitting of all flights regardless of transportation modes in the GeoLife and the Nokia MDC dataset. The green points refer to the flight length samples obtained from the GeoLife and the Nokia MDC dataset, while the solid red line represents the best fitted truncated power-law distribution according to Akaike weights.
Figure 3. Lognormal fit for single transportation mode in the Geolife dataset. (a-d) Flight distribution of all transportation modes (Car/Taxi/Bus, Walk/Run, Subway/Train, Bike). The solid blue line represents the best fitted lognormal distribution.
Figure 4. Flight length correlation for each transportation mode. (a-d) Flight length correlation of all transportation modes (Car/Taxi/Bus, Walk/Run, Subway/Train, Bike) in the GeoLife dataset. The Figure shows a high density of points near diagonal line $l_t = l_{t+1}$, identifying a small difference $l_{t+1} - l_t$ in the same transportation mode between two time steps.
| Dataset      | Transportation Mode | Fitted Distribution | p-value | Parameters |
|-------------|---------------------|--------------------|---------|------------|
| Overall     | Truncated Power-law | 0.00               | α = 1.57, λ = 0.00025 |
| Walk/Run    | Lognormal           | 0.00               | μ = 4.08, σ = 0.76   |
| GeoLife     | Bike                | Lognormal          | 0.00    | μ = 5.03, σ = 0.68   |
|             | Car/Bus/Taxi        | Lognormal          | 0.00    | μ = 5.78, σ = 1.04   |
|             | Subway/Train        | Lognormal          | 0.00    | μ = 7.27, σ = 0.51   |
| Overall     | Truncated Power-law | 0.00               | α = 1.39, λ = 0.00016 |
| Walk/Run    | Lognormal           | 0.00               | μ = 4.58, σ = 1.09   |
| Nokia MDC   | Bike                | Lognormal          | 0.00    | μ = 5.80, σ = 1.08   |
|             | Car/Bus/Taxi        | Lognormal          | 0.00    | μ = 6.89, σ = 0.91   |
|             | Subway/Train        | Lognormal          | 0.00    | μ = 6.93, σ = 0.94   |

TABLE 2. Parameters of fitted distributions in the GeoLife and in the Nokia MDC datasets. The p-value is less than 0.01 in all the fitted distributions, identifying a strong statistical significance.
SUPPLEMENTARY INFORMATION FOR
EXPLAINING THE POWER-LAW DISTRIBUTION OF HUMAN MOBILITY THROUGH
TRANSPORTATION MODALITY DECOMPOSITION

Kai Zhao, Mirco Musolesi, Pan Hui, Weixiong Rao, Sasu Tarkoma
Supplementary Figure S1. Lognormal fit for single transportation mode in the Nokia MDC dataset. (a-d) Flight distribution of all transportation modes (Car/Taxi/Bus, Walk/Run, Subway/Train, Bike). The solid blue line represents the best fitted lognormal distribution.
Supplementary Figure S2. Flight length correlation for each transportation mode. (a-d) Flight length correlation of all transportation modes (Car/Taxi/Bus, Walk/Run, Subway/Train, Bike) in the Nokia MDC dataset. The Figure shows a high density of points near diagonal line $l_t = l_{t+1}$, identifying a small difference $l_{t+1} - l_t$ in the same transportation mode between two time steps.
Supplementary Figure S3. Exponential elapsed time. The elapsed time $t$ is weighted exponentially between the different transportation modes. The exponentially weighted time interval is mainly due to a large portion of Walk/Run flight intervals.

Supplementary Figure S4. The power-law distribution of street length. This figure shows that the fitted road length distribution is very different to our truncated power-law fit in flights distribution. The flight length tails in the human mobility is much larger than those shown in this figure.
| Distribution            | Probability density function (pdf) |
|-------------------------|-----------------------------------|
| Truncated power-law     | $C x^{-\alpha} e^{-\lambda x}$   |
| Lognormal               | $\frac{1}{\sqrt{2\pi} \sigma x} \exp \left[ \frac{-(\ln(x) - \mu)^2}{2\sigma^2} \right]$ |
| Power-law               | $C x^{-\alpha}$                   |
| Exponential             | $\lambda e^{-\lambda x}$         |

Supplementary Table S1. Fitted distributions.

| Geolife     | Truncated Lognormal | Power-law | Exponential |
|-------------|---------------------|-----------|-------------|
| Overall     | 1.0000              | 0.0000    | 0.0000      | 0.0000      |
| Car/Bus/Taxi| 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Subway/Train| 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Walk/Run    | 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Bike        | 0.0000              | 1.0000    | 0.0000      | 0.0000      |

| Nokia MDC   | Truncated Lognormal | Power-law | Exponential |
|-------------|---------------------|-----------|-------------|
| Overall     | 1.0000              | 0.0000    | 0.0000      | 0.0000      |
| Car/Bus/Taxi| 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Subway/Train| 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Walk/Run    | 0.0000              | 1.0000    | 0.0000      | 0.0000      |
| Bike        | 0.0000              | 1.0000    | 0.0000      | 0.0000      |

Supplementary Table S2. Akaike weights of fitted distributions in the Geolife and the Nokia MDC datasets.

| Geolife     | Pearson correlation $r$ | $p$     |
|-------------|-------------------------|---------|
| Car/Bus/Taxi| 0.3640                  | 0.0000  |
| Subway/Train| 0.6445                  | 0.0000  |
| Walk/Run    | 0.5402                  | 0.0000  |
| Bike        | 0.5584                  | 0.0000  |

| Nokia MDC   | Pearson correlation $r$ | $p$     |
|-------------|-------------------------|---------|
| Car/Bus/Taxi| 0.3980                  | 0.0000  |
| Subway/Train| 0.4681                  | 0.0000  |
| Walk/Run    | 0.4570                  | 0.0000  |
| Bike        | 0.5291                  | 0.0000  |

Supplementary Table S3. Pearson correlation coefficient for consecutive flights length in the Geolife and the Nokia MDC dataset.
Supplementary Note 1

Given

$$P(x) = \int_{t=0}^{\infty} \frac{\lambda \exp(-\lambda t)}{x \sqrt{2\pi \sigma^2}} \exp[-\frac{(\ln(x) - \mu t)^2}{2\sigma^2 t}] \, dt.$$  \hspace{1cm} (11)

The calculation to obtain $\alpha'$ is as follows,

$$P(x) = \int_{t=0}^{\infty} \frac{\lambda \exp(-\lambda t)}{x \sqrt{2\pi \sigma^2}} \exp[-\frac{(\ln(x) - \mu t)^2}{2\sigma^2 t}] \, dt$$

$$= \frac{\lambda}{\sigma \sqrt{2\pi}} x^{-1} \int_{t=0}^{\infty} \exp(-\lambda t) \exp[-\frac{(-\ln(x) - \mu t)^2}{2\sigma^2 t}] \frac{1}{\sqrt{t}} \, dt$$

$$= \frac{\lambda}{\sigma \sqrt{2\pi}} x^{-1} \int_{t=0}^{\infty} \exp[-\frac{(\ln(x) - \mu t)^2}{2\sigma^2 t}] \frac{1}{\sqrt{t}} \, dt$$

$$= \frac{\lambda}{\sigma \sqrt{2\pi}} x^{-1} \exp(-\ln(x) \mu) \int_{t=0}^{\infty} \exp[-(\frac{\mu^2 + 2\lambda \sigma^2}{2\sigma^2})t - \frac{(\ln(x))^2}{2\sigma^2 t}] \frac{1}{\sqrt{t}} \, dt.$$  

Using the substitution $t = u^2$ gives

$$P(x) = \frac{\lambda}{\sigma \sqrt{2\pi}} x^{-1} \exp(-\ln(x) \mu) \int_{u=0}^{\infty} \exp[-(\frac{\mu^2 + 2\lambda \sigma^2}{2\sigma^2})u^2 - \frac{(\ln(x))^2}{2\sigma^2} \frac{1}{u^2}] \frac{1}{\sqrt{u^2}} \, 2udu.$$  

Let $a = \frac{\mu^2 + 2\lambda \sigma^2}{2\sigma^2}$ and $b = (\ln(x))^2(2\sigma^2)$, from the integral table we get

$$\int_{u=0}^{\infty} \exp(-au^2 - \frac{b}{u^2}) = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp(-2\sqrt{ab}),$$

which helps us to get the expression for $P(x)$,

$$P(x) = \frac{\lambda}{\sigma \sqrt{\frac{\mu^2}{\sigma^2} - 2\lambda^2}} x^{-1} \left( \frac{\sqrt{\mu^2 + 2\lambda \sigma^2}}{\sigma^2} \right)$$

$$= \frac{\lambda}{\sigma \sqrt{\frac{\mu^2}{\sigma^2} - 2\lambda^2}} x^{-\alpha'}.$$

Finally the expression for $\alpha'$

$$\alpha' = 1 - \frac{\mu}{\sigma^2} + \frac{\sqrt{\mu^2 + 2\lambda \sigma^2}}{\sigma^2}.$$