Performance of Multi-path Routing in MANET with Long-Tailed Traffic

Zafar Q. Taha and Xian Liu
Department of Systems Engineering
University of Arkansas at Little Rock
2801, S University Ave, Little Rock, AR, USA
{zqtaha,xxliu}@ualr.edu

Abstract

The dynamic source routing (DSR) method has received much attention in the studies and practice of mobile ad hoc networking. Efforts toward mitigating the overhead have been made by using the multi-path routing strategy. Conventional performance evaluations were conducted with the exponential distribution of the link life. In this paper, we evaluate the performance through a long-tailed distribution called the compound log-normal distribution. The reliability archetypes of the series and parallel connections have been used in developing the concerned models. The overall system reliability of ad hoc networks with multi-path routing is also investigated.

1. Introduction

Interest in the mobile ad hoc network (MANET) is high. Besides its indispensable role in the battlefield, the MANET’s civilian applications include the intelligent device networking, the mobile robot communication, the on-the-fly conferencing, the sensor networking, and the home-area wireless networking [5]. Technically, a typical MANET is characterized by the marginal power resources and stochastic mobility of the nodes. Besides, since the transmission range of a single wireless hop is physically restricted, in an MANET, most communications are conducted by using multiple wireless hops across the concerned region. In principle, routing for the multi-hop paradigm can be accomplished through either a single path or multiple (parallel or disjoint) paths. In general, the former is less flexible than the latter, regarding the traffic control, load balance, reliability, and performance. In the context of multi-hop-multi-path routing, much attention has paid to the on-demand protocols. In this area, one of the main concerns is the routing overhead, usually measured by the number of the control packets transmitted, as opposed to the user or pay-load packets. It has been recognized that the routing overhead of the on-demand protocols is most likely lower than the shortest path protocols. The dynamic source routing (DSR) method is one of the representative on-demand protocols.

The DSR works in two main phases: path discovery and path maintenance. The path discovery phase is initiated by the source node. A query message is sent to the whole network (a.k.a. flooding) for the purpose of finding a path to the destination. Each query message carries a sequence of hops it traversed in its message header. Once the query reaches the destination, the destination initiates a reply packet back to the source. The reply packet contains a list of nodes along the path from the source to the destination. In the path discovery phase, more than one path may be discovered and only the disjoint paths are stored in the source’s path cache. The path maintenance phase ensures that the paths stored in the path cache are valid. If the data link layer of a node detects a transmission error, then the node creates a path error packet and transmits it to the source. Upon receiving the path error packet, the sources check their path cache and delete the paths containing the failed links. Then they can either try to use other alternate paths in their caches or invoke a new path discovery phase.

The main performance bottleneck in DSR is flooding. In most cases, flooding takes up a substantial amount of network bandwidth. There is a strong need to reduce the frequency of flooding for path discovery. A variety of methods have been reported in the literature. However, the performances of these methods were almost exclusively evaluated by the exponential models [4]. Though this approach might be appropriate for the voice traffic, it is not suitable for the data traffic. For the latter, studies have shown that the long-tailed distributions are more adequate. In the present work, we use the log-normal distribution, one of the most representative long-tailed distributions, to conduct the analysis. The main objective is to investigate the situation where the long-tailed data traffic is dominant.

The rest of this paper is organized as follows: The motivation and rational to adopt a long-tailed distribution are
addressed in section 2. Then the mathematical essentials are developed in Section 3. Next, the performance analysis of the multi-path protocol is compared with the single path protocol in Section 4, where the numerical results are included to complement the analysis. In Section 5, the system availability is discussed and several scenarios are compared. Finally, the conclusions are given in Section 6.

2. Background and Motivation

Ideally we would like to have analytic models for every process: simple mathematical descriptions rather than huge statistical data. When modeling network traffic, connection duration are often assumed to be Poisson processes because such processes have attractive theoretical properties[6]. The inadequacy of exponential distribution for modeling connection duration is shown in [1] and log-normal distribution is suggested as a better choice. Recent work argues convincingly that data traffic connection duration is much better modeled using statistically self-similar(SSM) processes[3]. One of the primary attributes of the SSM traffic is the long tailed distributions such as log-normal distribution.

A wireless link in ad hoc network can fail due to many reasons, we assume that the link lifetime is mainly affected by the point-to-point communication duration. Accordingly, the lifetime of a wireless link can be described by a random variable. Our work follows the trend in [4] to investigate the performance of the multi-path protocol based on long tailed distribution - log-normal distribution. A random variable is said to follow log-normal distribution if it’s probability density function (PDF) and cumulative distribution function (CDF) respectively are:

\[
 f_X(x) = \begin{cases} 
 \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[ -\frac{(\ln x - \mu)^2}{2\sigma^2} \right] & x > 0 \\
 0 & x \leq 0 
\end{cases} \quad (1)
\]

\[
 F_X(x) = \Phi \left( \frac{\ln x - \mu}{\sigma} \right) \quad (2)
\]

where \(-\infty < \mu < \infty, \sigma > 0\) and

\[
 \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-y^2/2)dy
\]

where \(\mu\) and \(\sigma\) are mean and variance of normal distribution \(N(\mu, \sigma)\) respectively. Accordingly, the mean, and variance respectively are:

\[
 E(x) = \exp(\mu + \frac{\sigma^2}{2}) \quad (3)
\]

\[
 \text{var}(x) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2) \quad (4)
\]

3. Essentials of Compound Log-normal distribution

In MANET, a path is defined as a series of \(n\) links. Thus a path fails when any one of the links fail. We represent the lifetime of wireless link between two nodes by a random variable. The attention is then paid to a particular path \(P\) formed by a pair of source-destination (SD) nodes connected by a series of \(n\) links. Let us denote each link with \(L_i\) and its lifetime with with \(X_{Li}\) with \((i = 1, 2, ..., n)\). As a result, the lifetime of a path \(P\) with links can be expressed as:

\[
 X_P = \min(X_{L_1}, X_{L_2}, ..., X_{L_n}) \quad (5)
\]

Accordingly, reliability function of path \(P\) with \(n\) series connections is [8]:

\[
 R_{ss}(x) = \prod_{i=1}^{n} R_{L_i}(x) \quad (6)
\]

\[
 1 - F_{X_P}(x) = 1 - \prod_{i=1}^{n} (1 - F_{X_{L_i}}(x)) \quad (7)
\]

\[
 F_{X_P}(x) = 1 - (1 - F_{X_{L_i}}(x))^n \quad (8)
\]

where \(R_X(x)\) is the reliability function, defined as \(R_X = 1 - F_X(x)\) and subscript ”ss” represents series connection. Equation [9] is obtained assuming homogeneous nature of links, and taking derivative on both sides we get the PDF:

\[
 f_{X_P}(x) = n[1 - F_X(x)]^{n-1} f_X(x) \quad (9)
\]

The validity of above functions is shown below. From the basic properties of CDF, we have:

\[
 F_{X_P}(0) = 1 - [1 - F_X(0)]^n = 1 - 1 = 0
\]

\[
 F_{X_P}(\infty) = 1 - [1 - F_X(\infty)]^n = 1 - (1 - 1)^n = 1
\]

From the PDF properties we have:

\[
 \int_{x=0}^{\infty} f_{X_P}(x)dx = \int_{0}^{\infty} n[1 - F_X(x)]^{n-1} f_X(x)dx
\]

Let \(u = 1 - F_X(x)\), therefore \(du = -f_X(x)dx\) substituting these equations, we get

\[
 \int_{0}^{1} nu^{n-1}du = u^{n}_{|0} = 1
\]

Thus, distributions given by equations (7) and (8) are valid distribution functions and we will address them as compound Log-normal distribution functions.

Expressions for mean and median are as follows:

\[
 E_{ss}(X_P) = \int_{x=0}^{\infty} n[1 - F_X(x)]^{n-1} f_X(x)dx \quad (10)
\]
Figure 1. CDF of Compound Log-normal distribution

$B_s(X) = 1 - [1 - F_X(x)]^n = 0.5$ \hspace{1cm} (11)

These expressions cannot be evaluated analytically due to presence of $\exp(-x^2)$ in the integral; hence we resort to numerical methods to evaluate these integrals. Note subscript "s" corresponds to single-path paradigm. Table I shows averages for $\mu = 0$, $\sigma = 1$ For multi-path case, assume there are distinct paths each with length $n_q$, $(q = 1, 2, ..., m)$ for a given SD pair. A new path discovery is needed only after all paths broke. These multiple paths that exists could be considered as parallel connections to same SD pair, thus time between successive path discoveries can be expressed as:

$Z = \text{min}(X_1, X_2, ..., X_n)$ \hspace{1cm} (12)

Accordingly, the reliability function of $Z$ is given by [8]:

$R_{pl}(Z) = 1 - \prod_{p=1}^{m} [1 - R_{X_p}(x)]$ \hspace{1cm} (13)

Table 1. Single path averages for $n$ links

<table>
<thead>
<tr>
<th>$n$</th>
<th>Mean</th>
<th>Median</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.648</td>
<td>1.000</td>
<td>4.642</td>
</tr>
<tr>
<td>2</td>
<td>0.990</td>
<td>0.579</td>
<td>1.325</td>
</tr>
<tr>
<td>3</td>
<td>0.500</td>
<td>0.441</td>
<td>0.433</td>
</tr>
<tr>
<td>4</td>
<td>0.284</td>
<td>0.368</td>
<td>0.228</td>
</tr>
</tbody>
</table>

Figure 2. PDF of Compound Log-normal distribution

4. Performance analysis of the multi-path routing

Let us consider a case where all paths are to have same length $n$ for simplicity of analysis. We have from eqn (12):

$F_Z(z) = \prod_{p=1}^{m} F_{X_p}(z)$ \hspace{1cm} (14)

where subscript "pl" represents parallel multi-paths. The PDF may be obtained by differentiating on both sides:

$f_Z(z) = \frac{d}{dz} F_Z(z) = \sum_{q=1}^{m} \frac{F_{X_q}}{F_{X_q}} \prod_{p=1}^{m} F_{X_p}
= \prod_{p=1}^{m} F_{X_p} \left[ \sum_{q=1}^{m} q^{m} \frac{f_{X_q}}{F_{X_q}} \right]
= \prod_{p=1}^{m} [1 - [1 - F_X(x)]^{n_q}]
\left[ \sum_{q=1}^{m} n_q [1 - F_X(x)]^{n_q-1} f_X(x) \right]$ \hspace{1cm} (15)
The CDF of compound log-normal distribution for multi-path is shown in Figure 1. Note that distribution is dependent on \( n \), the path length. The validity of expressions [14] and [15] can be obtained by following similar approach used to show validity of distributions given by equations [7] and [8]. The PDF with \( n = 4 \), and \( m = 1, 2, 4, 10 \) are illustrated in Figure 2. Similar expressions for mean and median can be written for the multi-path case:

\[
E_m(Z) = \int_{z=0}^{\infty} zm \{1 - [1 - F_X(x)]^n \}^{m-1} \left[1 - F_X(x)\right]^{n-1} f_X(x) dx
\]

\( B_m(Z) = \{1 - [1 - F_{X_m}(z)]^n \}^m = 0.5 \) \hspace{1cm} (19)

where the subscript ”\( m \)” corresponds to m-path paradigm.

Table 2 shows the values of mean, median and variance for \( n = 4, \mu = 0, \sigma = 1 \) for multi-path case.

### Table 2. Multi path averages for \( m \) paths

<table>
<thead>
<tr>
<th>( m )</th>
<th>Mean</th>
<th>Median</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.910</td>
<td>0.368</td>
<td>5.217</td>
</tr>
<tr>
<td>2</td>
<td>0.192</td>
<td>0.533</td>
<td>0.381</td>
</tr>
<tr>
<td>3</td>
<td>0.085</td>
<td>0.715</td>
<td>0.106</td>
</tr>
<tr>
<td>4</td>
<td>0.050</td>
<td>0.912</td>
<td>0.047</td>
</tr>
</tbody>
</table>

In order to estimate the relative merit of the multi-path approach, we evaluate the ratio of \( E_m(Z) \) to \( E_1(Z) \). The ratio \( h(n, m) = E_m(Z)/E_1(Z) \) represents the normalized average time between successive path discoveries. Recall that a new path discovery is initiated only when all the existing paths to the destinations have been broken. Therefore, the larger values of \( h(n, m) \) should be sought whenever possible. The numerical profile of three instances for \( m = 2, 3, \) and \( 5 \) are illustrated in Figure 3.

The merit of employing paths with the identical length can be easily observed from Figure 3. However, the advantage tends to diminish when parameter \( n \) gets larger.

Since the median is used more often in the field, it may be desirable to estimate the relative merit by similar ratio as \( h \). The table III shows the value of median for different values of \( n \) and \( m \). This table is used to develop Figure 4 which has ratio of medians of m-path to 1-path for varying values of \( n \).

### Table 3. Median ratio m-path to 1-path

<table>
<thead>
<tr>
<th>( n )</th>
<th>( m = 1 )</th>
<th>( m = 2 )</th>
<th>( m = 3 )</th>
<th>( m = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
<td>0.579</td>
<td>0.441</td>
<td>0.368</td>
</tr>
<tr>
<td>2</td>
<td>1.724</td>
<td>0.901</td>
<td>0.654</td>
<td>0.532</td>
</tr>
<tr>
<td>3</td>
<td>2.269</td>
<td>1.122</td>
<td>0.794</td>
<td>0.637</td>
</tr>
<tr>
<td>4</td>
<td>2.713</td>
<td>1.292</td>
<td>0.900</td>
<td>0.714</td>
</tr>
</tbody>
</table>

5. Reliability of ad hoc network with multiple paths versus single path

In MANET, reliability of paths is dependent upon the individual links working state. In this section we show that multi-path is more reliable than single path protocol. Our analysis is conducted under the framework developed in ([8], chapter 9, section 7). Assume that all links are working initially and let

\[ A(t) = P \{ \text{system is working at time } t \} \]

\( A(t) \) is availability of system at time \( t \). Assuming that the ON and OFF distribution for each link \( i \) in the path follow continuous Log-normal distribution with respective means \( 1/\lambda_i \), and \( 1/\mu_i \), then from it follows that

\[ A_i(t) = \frac{1/\lambda_i}{1/\lambda_i + 1/\mu_i} = \frac{\mu_i}{\mu_i + \lambda_i} \]

As \( t \to \infty \), then the limiting availability- call it \( A \) is

\[ A = \lim_{t \to \infty} A(t) = R \left( \frac{\mu}{\mu + \lambda} \right) \]

where \( A \) depends only on the ON and OFF distributions through their means. And \( R(\cdot) \) is reliability function.

For a single path with \( n \) homogeneous nodes connected in series, we have:

\[ A_s = \prod_{i=1}^{n} \frac{\mu_i}{\mu_i + \lambda_i} = \left( \frac{\mu}{\mu + \lambda} \right)^n \] \hspace{1cm} (20)
For $m$ multiple parallel path each with $n$ nodes, we have:

$$A_m = 1 - \prod_{i=1}^{m} \left( 1 - \left( \frac{\mu_i}{\mu_i + \lambda_i} \right)^n \right) = 1 - \left[ 1 - \left( \frac{\mu}{\mu + \lambda} \right)^n \right]^m$$

(21)

In order to evaluate the merit of the multi-path approach, we take the ratio $I(n, m) = A_m / A_s$. Figure 5 shows that the multi-path protocol is more reliable in MANET than the single path protocol. Reliability here is interpreted in terms of availability of paths in MANET.

6. Conclusion

In this paper we evaluate the efficiency of ad hoc network in terms of reduced number of path discoveries in the MANET with source initiated routing method. This model is based on long-tailed traffic distribution namely Log-normal distribution. The comparison between multi-path approaches to single-path is given. Although the former is generally better than the latter, it is not necessary to introduce a great number of secondary paths. The reliability of the network as a whole is considered and is shown that the multi-path protocol provides more reliable communication paths than single path protocol. The above conclusions extend the results reported in the early studies to the bursty data traffic paradigms since our model takes the long-tailed behaviors into account.

References