A Prediction-based Link Availability Estimation for Mobile Ad Hoc Networks

Shengming Jiang[†], Dajiang He[‡] and Jianqiang Rao[‡]
[†]Centre for Wireless Communications
jiangsm@ieee.org
‡ Department of Electrical and Computer Engineering
{engp8587,engp9040}@nus.edu.sg
National University of Singapore

Abstract--- **One critical issue for routing in mobile ad hoc networks (MANET) is how to select a reliable path that can last longer since mobility may cause radio links to break frequently. To answer this question, a criterion that can judge path reliability is needed. The reliability of a path depends on the availability of the links[1] constituting the path. However, how to measure link availability to answer the above question has not been addressed adequately in the literature. In this paper, a prediction-based link availability estimation is introduced and verified through computer simulations. This estimation algorithm can be used to develop a metric for path selection in terms of path reliability, which can improve the network performance as to be shown by the simulation results.**

Keywords--- **MANET, routing metric, path reliability, link availability and mobility.**

I. Introduction

Routing is difficult in MANET since mobility may cause radio links to break frequently. When any link of a path breaks, this path needs to be either repaired by finding another link if any or replaced with a newly found path. This rerouting operation costs the scarce radio resource and battery power while rerouting delay may affect quality of service (QoS) for applications and degrade the network performance. To reduce rerouting operation, selecting an optimal path in such networks should consider path reliability more than some metrics used in wired networks such as path cost and QoS etc. The reliability of a path depends on the availability of all links constituting this path. However, most routing schemes in the literature focus mainly on the procedure of information exchange for finding and/or maintaining a path between two nodes, and often use 'shortest path' (measured in terms of the number of hops or links that a path goes through) as the major routing metric [1]. How to measure link availability properly in order to quantify a routing metric in terms of path reliability has not been addressed adequately.

In [2] and [3], a so-called 'associativity' is defined as a new routing metric for link reliability. This metric tries to reflect the degree of the association stability between two mobile nodes through the connection stability of a node with respect to another one over time and space. Each node generates a beacon to signify its existence periodically. Upon receiving a beacon, the receiver increases the value of its associativity with the beaconing node. In [4], both signal stability and location stability are used to quantify the reliability of a link. With the signal stability, each node classifies its neighbors as either 'strongly connected' or 'weakly connected' according to the signal strength of received beacons generated periodically by its neighbors. The location stability is measured in terms of the period of time that a link has existed. Accordingly, the routing metric biases the selected path toward the one consisting of strong channels which have been in existence for a time greater than some threshold. A common weakness of the above two pure measurement-based criteria for link reliability is that they cannot reflect possible changes in link status happening in the future. That is, the reliability of a link measured as 'better' based on past and/or current information on link status may become worse with time than that of those currently measured as 'worse' due to the dynamic nature of mobile environments. This possible misjudgment to link reliability would affect the network performance especially in a high mobility environment.

A probabilistic link availability model which can predict the future status of a wireless link is proposed in [5] and [6]. In this model, the link availability is defined as the probability that there is an active link between two mobile nodes at time $t_0 + T$ ($T > 0$) given that there is an active link between them at time t_0. Note that a link is still considered available at $t_0 + T$ even if it experienced failures during one or more intervals between t_0 and $t_0 + T$. So, this link availability can be viewed as $\frac{T_a}{T}$, where T_a is the sum of all non-continuous time periods that the link is available between t_0 and $t_0 + T$. This metric can be used by a node to select more reliable neighbors to form a stabler cluster, but it is not practical to use it for path selection as explained below. As mentioned earlier, any link of a path breaks, a rerouting is required immediately. It is unlikely to let the related nodes to wait for this broken link to become available again. With this consideration, it is more practical to use the continuous time period (T_c) that a link will last from time t_0 for link selection performed at this moment. However, a longer T_a does not always mean a longer T_c and vice versa. Another weakness of this model is that it does not make use of some information that can be measured in order to predict more precisely link availability for short time periods. As shown by the results in [5], this model can match the simulation results well when the predicted time periods longer than tens minutes. However, it substantially under-estimates the link availability for time periods shorter than several minutes, which are actually

[1] By 'a link being available', we mean that the radio quality of the link satisfies the minimal requirement of the successful communication. 'Link availability' is a general term to measure the probability or degree that a link is in the above available state. Words 'availability' and 'reliability' will be used interchangeable to describe link status in this paper.

most interesting to routing since typical flow duration is often less than several minutes. For example, the average TCP flow duration is about 12 $\sim$ 19 seconds on international Internet links with USA [7].

In this paper, we introduce a prediction-based link availability estimation as well as a routing metric in terms of path reliability based on this estimation. The basic idea of this estimation is to let a node to first predict a continuous time period (T_p) that an currently available link will last from time t_0 by assuming that both nodes of the link will keep their current movements (i.e., speed and direction) unchanged. Then, we try to estimate the probability that the link will last to $t_0 + T_p$, $L(T_p)$, by considering possible changes in the nodes' movements occurring between t_0 and $t_0 + T_p$. More precisely, the link availability estimation consists of 'unaffected T_p' with the above assumption being held and 'affected T_p' with movement being changed. As discussed later, it is difficult to give an accurate calculation of $L(T_p)$. However, we think a reasonable estimation of $L(T_p)$ can be still helpful for link selection in terms of reliability.

Regarding T_p prediction, a measurement-based scheme has been proposed in [8] and [9], in which, a node can predict T_p for an active link with another node by measuring relative distances between them without knowing the velocities of their movements. Recently in [10], a similar scheme has been also proposed to predict T_p, in which, the velocity of a node's movement is supposed to be known by using Global Position Systems (GPS). Both T_p predictions done at t_0 assume that the related nodes' movements between t_0 and $t_0 + T_p$ will be the same as these observed at t_0 through measurement or GPS. Due to the space limitation, we will focus only on the estimation of $L(T_p)$ and a routing metric based on $L(T_p) \times T_p$ in this paper. The readers please refer to [8], [9] and [10] for more details on T_p prediction.

The paper is organized as follows. Section II gives a detailed description of the proposed link availability estimation. The results given by this estimation are compared with simulation results in Section III. A routing metric in terms of path reliability based on the estimation is discussed in Section IV. Section V concludes the paper.

II. A LINK AVAILABILITY ESTIMATION

The basic assumptions for the proposed estimation algorithm are similar to those used in the literature such as [5] and [6], that is,

- Mobility epoch[2] lengths are exponentially distributed with mean λ^{-1}, i.e.,

$$\begin{aligned} E(x) &\stackrel{\triangle}{=} P\{\text{Epoch length} \leq x\} \\ &= 1 - e^{-\lambda x}. \end{aligned}$$

- Node mobility is uncorrelated.

To simplify the discussion, we further assume that each node has the same mean epoch length (i.e., λ^{-1}). However, the following derivation can be extended for the case of different mean epoch lengths.

Given a prediction T_p on the continuously available time for an active link between two nodes at time t_0, the availability of this link, $L(T_p)$, is defined as

$$L(T_p) \stackrel{\triangle}{=} P\{\text{To last to } t_0 + T_p \mid \text{Available at } t_0\},$$

which indicates the probability that the link will be continuously available from time t_0 to $t_0 + T_p$.

The calculation of $L(T_p)$ can be divided into two parts: the link availability when the velocities of the two nodes keep unchanged between t_0 and $t_0 + T_p$, $L_1(T_p)$, and the one for the other cases, $L_2(T_p)$. That is,

$$L(T_p) = L_1(T_p) + L_2(T_p). \tag{1}$$

It is easy to calculate $L_1(T_p)$, which is equal to the probability that the epochs from t_0 onwards for the two nodes are longer than T_p because T_p is an accurate prediction if the movements of the two nodes keep unchanged [8], [9] and [10]. Since nodes' movements are independent of each other and exponential distribution is 'memoryless' [11], $L_1(T_p)$ is given by

$$\begin{aligned} L_1(T_p) &= [1 - E(T_p)]^2 \\ &= e^{-2\lambda T_p}. \end{aligned} \tag{2}$$

However, it is difficult to give an accurate calculation for $L_2(T_p)$ because of the difficulties in learning changes in link status caused by changes in a node's movement. For example, what is the probability for a link to be continuously available after a movement change happens and how many changes in movement will happen between t_0 and $t_0 + T_p$ and etc? In the following, we discuss an estimation for $L_2(T_p)$.

Denote $\Phi < T_p$ as a random variable for the time interval between t_0 and $t_0 + T_p$ during which either of the two nodes or both change their movements. $P\{\phi \leq \Phi < T_p\}$ indicates the probability for both nodes to keep their movements unchanged between t_0 and $t_0 + \phi$ while either of them or both change after $t_0 + \phi$. The calculation of $P\{\phi \leq \Phi < T_p\}$ is divided for two cases: when only one node changes its movement and when both nodes change their movements between $t_0 + \phi$ and $t_0 + T_p$. From the same reason as for deriving Eq. 2, $P\{\phi \leq \Phi < T_p\}$ is given by

$$\begin{aligned} P\{\phi \leq \Phi < T_p\} &= 2[E(T_p) - E(\phi)][1 - E(T_p)] \\ &\quad + [E(T_p) - E(\phi)]^2 \\ &= e^{-2\lambda\phi} - e^{-2\lambda T_p}. \end{aligned} \tag{3}$$

$\mathcal{L}_2(\phi)$ is further introduced to estimate the link availability corresponding to ϕ as follows:

$$\mathcal{L}_2(\phi) = \frac{\phi + (T_p - \phi)pe^{-2\lambda(T_p - \phi)}}{T_p} + \epsilon, \tag{4}$$

where p is the probability for the two nodes to move to closer each other after changing their movements and $\epsilon \geq 0$ is an adjustment to the link availability calculated by the first part on the right hand of Eq. 4 (i.e., $\frac{[.]}{T_p}$). Both are supposed to be independent of ϕ. More discussion on p and ϵ is given later.

[2] An epoch is a random length interval during which a node moves in a constant direction at a constant speed.

Eq. 4 tries to estimate link availability through calculating the total time (T_t) that a link will be continuously available between t_0 and $t_0 + T_p$ if movement changes happen as discussed below. After the first movement change happens at $t_0 + \phi$, the link between the related nodes can be still continuously available if this change makes these two nodes to move close each other with probability p. The link is expected to be continuously available from $t_0 + \phi$ to $t_0 + T_p$ if the two nodes keep their movements unchanged during this period with probability $e^{-2\lambda(T_p-\phi)}$ (the derivation is similar to Eq. 2). So, T_t can be calculated by

$$T_t \;=\; \phi + (T_p - \phi)pe^{-2\lambda(T_p-\phi)} + ... \tag{5}$$

The '...' part in Eq. 5 means the accurate T_t can be achieved by repeating the above calculation with considering the second possible changes in movements during the remaining period and so on. However, doing so will complicate the calculation. On the other hand, it is reasonable to assume that the averaged contribution of this part (i.e., '...') to the overall link availability is relatively smaller than the sum of $L_1(T_p)$ and $\frac{\phi+(T_p-\phi)pe^{-2\lambda(T_p-\phi)}}{T_p}$. For the sake of simplicity, ϵ is introduced to estimate the link availability contributed by this part. As discussed later, the use of measured ϵ can also make the proposed estimation adaptive to environmental changes.

The average $\mathcal{L}_2(\phi)$ over ϕ, $\overline{\mathcal{L}}_2$, is used to estimate $L_2(T_p)$. $\overline{\mathcal{L}}_2$ is given by

$$\overline{\mathcal{L}}_2 \;=\; \int_0^{T_p} \mathcal{L}_2(\phi) f(\phi) d\phi, \tag{6}$$

where $f(\phi) \geq 0$ is given by

$$\begin{aligned} f(\phi) &= \lim_{\triangle\phi \to 0} \frac{P\{\phi \leq \Phi < T_p\} - P\{\phi + \triangle\phi \leq \Phi < T_p\}}{\triangle\phi} \\ &= -\frac{dP\{\phi \leq \Phi < T_p\}}{d\phi} \\ &= 2\lambda e^{-2\lambda\phi}. \end{aligned} \tag{7}$$

Substitute $\mathcal{L}_2(\phi)$ and $f(\phi)$ in Eq. 6 with Eqs. 4 and 7, respectively. Then $\overline{\mathcal{L}}_2$ can be estimated by[3]

$$\begin{aligned} \overline{\mathcal{L}}_2 &\approx \int_0^{T_p} \Big(\frac{\phi + (T_p - \phi)pe^{-2\lambda(T_p-\phi)}}{T_p} + \epsilon\Big) 2\lambda e^{-2\lambda\phi} d\phi \\ &= \frac{2\lambda}{T_p} \int_0^{T_p} [\phi e^{-2\lambda\phi} + pe^{-2\lambda T_p}(T_p - \phi) + \epsilon T_p e^{-2\lambda\phi}] d\phi \\ &= \frac{1}{2\lambda T_p} + \epsilon + e^{-2\lambda T_p}(p\lambda T_p - \frac{1}{2\lambda T_p} - \epsilon - 1). \end{aligned} \tag{8}$$

So, we have an estimation of $L(T_p)$ as follows:

$$\begin{aligned} L(T_p) &\approx L_1(T_p) + \overline{\mathcal{L}}_2 \\ &= \frac{1}{2\lambda T_p} + \epsilon \\ &\quad + e^{-2\lambda T_p}(p\lambda T_p - \frac{1}{2\lambda T_p} - \epsilon). \end{aligned} \tag{9}$$

Now we discuss how to obtain p and ϵ. It is easy to get p for a free environment (where a node has no predefined movement trace) with following consideration. After a change in the movements of two nodes, they will move either to close each other or far away from each other. Since there is no particular factor to affect a node's movement, the above two events should happen in an equal probability such that $p = 0.5$. In this case, Eq. 9 can be further simplified into

$$\begin{aligned} L(T_p) &\approx (1 - e^{-2\lambda T_p})(\frac{1}{2\lambda T_p} + \epsilon) \\ &\quad + \frac{\lambda T_p e^{-2\lambda T_p}}{2}. \end{aligned} \tag{10}$$

The value of ϵ mainly depends on environmental factors such as node density and node's radio coverage and etc. Environments are changed as time and movement. It is impossible to give a mathematical ϵ that can satisfy all environments. With this respect, it is practical, from the point of view of application, to devise a measurement method for ϵ so that the proposed estimation algorithm can adapt environmental changes. Such a measurement is proposed below.

After a node has a prediction T_p on an active link with another node at time t_0, it then measures how long this link will really last from t_0, say T_r. If this link is still available just after $t_0 + T_p$, the node sets $T_r = T_p$ and does another prediction. Repeating the above operation, the node may have a series of 2-tuples $< T_{r,i}, N_i >$ for the same T_p value, where N_i is the occurrence times of the same T_r value (namely $T_{r,i}$). Then we can have a measured $L(T_p)$ corresponding to this T_p value, $L_m(T_p)$, which is calculated by

$$L_m(T_p) = \frac{\sum_{i=1}^{N_{T_r}} (\frac{T_{r,i}}{T_p} \times N_i)}{\sum_{i=1}^{N_{T_r}} N_i}, \tag{11}$$

where N_{T_r} is the total number of different T_r values observed. Substituting $L(T_p)$ in Eq. 10 with $L_m(T_p)$, we have a measured ϵ corresponding to this T_p value (ϵ_m) as

$$\epsilon_m \;\approx\; \frac{L_m(T_p) - \frac{\lambda T_p e^{-2\lambda T_p}}{2}}{1 - e^{-2\lambda T_p}} - \frac{1}{2\lambda T_p}. \tag{12}$$

Repeating the above operation and calculation for different T_p values, say $T_{p,j}$ ($j = 1, 2, ..$), we can have a series of ϵ_m, $\epsilon_{m,j}$. Then the final ϵ can be estimated by

$$\epsilon \;\approx\; \frac{\sum_{j=1}^{N_{T_p}} M_j \epsilon_{m,j}}{\sum_{j=1}^{N_{T_p}} M_j}. \tag{13}$$

where M_j is the occurrence times of $T_{p,j}$ and N_{T_p} is the total number of different T_p values predicted. At the beginning of operation, a node can set $\epsilon = 0$, with which Eq. 10 gives a conservative prediction for the link availability, $L_{min}(T_p)$. That is,

$$L_{min}(T_p) \;=\; \frac{1 - e^{-2\lambda T_p}}{2\lambda T_p} + \frac{\lambda T_p e^{-2\lambda T_p}}{2}. \tag{14}$$

[3] $\int xe^{ax}dx = \frac{e^{ax}(ax-1)}{a^2}$ [12].

III. Numerical results

The environment under simulation is a two dimensional space, in which there are 30 nodes moving randomly according to some mobility models. The maximal radius of a mobile's radio coverage is 300 meters which is a design parameter of some typical IEEE 802.11 products such as NOKIA A020 [14]. The transmission rate of a node is 2 $Mbit/s$ and IEEE 802.11 is adopted for the MAC layer. Mobility models with both exponentially and non-exponentially distributed epoch lengths have been investigated. The prediction method proposed in [8], [9] is adopted for a node to predict T_p. The simulation package is OPNET [15].

The simulation results are plotted with x-axis being the converged T_p and y-axis being the converged $L_m(T_p)$ given by Eq. 11 because there are too many points observed in the simulation. The convergence is done as follows. Suppose there are $J > 0$ triplets $< T_{p,j}, L_m(T_{p,j}), N_j >$ in which $I \leq T_{p,j} < I + 1$ (I is an integer and $j = 1, ...J$), where N_j is the occurrence times of $T_{p,j}$. This series is converged into one point as

$$x = \frac{\sum_{j=1}^{J} T_{p,j} N_j}{\sum_{j=1}^{J} N_j} \text{ and } y = \frac{\sum_{j=1}^{J} L_m(T_{p,j}) N_j}{\sum_{j=1}^{J} N_j}.$$

Then, x and y are used as T_p and $L_m(T_p)$ in Eq. 12 to calculate ϵ_m. The final ϵ is obtained by using Eq. 13 for all the merged points. This ϵ is used by Eq. 10 to plot $L(T_p)$. Since the observed T_p ranges from 0 to hundreds even thousands which depends on λ^{-1}, the above convergence can show more clearly the tendency of simulation results without losing the accuracy.

We first look at some results for the random walk-based mobility model [6]. In this model, a node first selects a direction uniformly from a given space and a speed uniformly from 0 to 20 m/s for its next movement. Then the length of the epoch for the above movement is selected according to an exponential distribution with mean λ^{-1}. With this model, a node is allowed to move beyond the boundary of the given space. But the direction of its next movement following a 'moving-out' is forced to the given space in order to maintain node density.

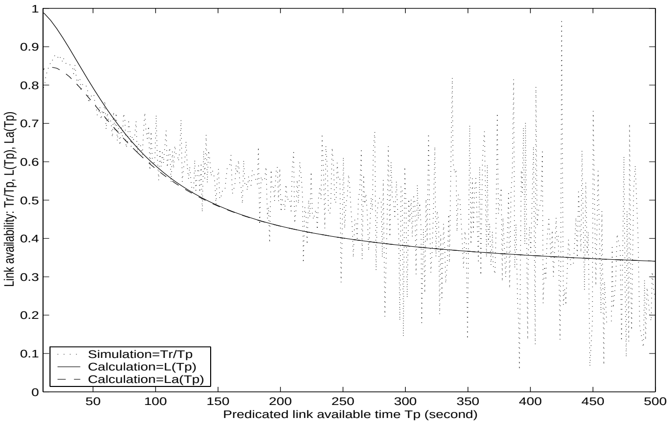
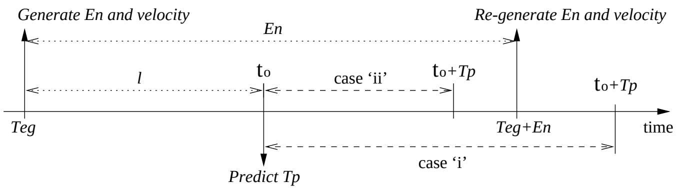


Fig. 1. Random walk-based model: $\lambda^{-1} = 60s$ and $(500m)^2$

Fig. 1 plots simulation results (dotted line) and calculation results given by Eq. 10 (solid line) for a space of $500 \times 500m^2$ and $\lambda^{-1} = 60s$. We can find that simulation results almost fluctuate around the curve given by $L(T_p)$. This means $L(T_p)$ can approximate $\frac{\overline{T_r}}{T_p}$, where $\overline{T_r}$ is the mean time that a link will be continuously available corresponding to a prediction T_p. However, there is a substantial mismatch between them when T_p is very short, i.e., a 'quasi-bell shape' appearing in the simulation results but not in the calculation results. This phenomenon is mainly caused by an intrinsic weakness of simulation for the memoryless property of exponential distribution as explained below.

As illustrated in Fig. 2 (on the next page), in the simulation, a node randomly generates an epoch length (E_n) as well as a velocity for its next movement at time T_{eg}. To save simulation time and simplify simulation, the above operation will be only repeated at $T_{eg} + E_n$ and so on. Therefore, the generated E_n will keep unchanged until $T_{eg} + E_n$. On the other hand, T_p prediction is independent of the above E_n generation. Suppose T_p is predicted at time t_0 and $T_{eg} < t_0 \leq T_{eg} + E_n$. As illustrated in Fig. 2, $t_0 + T_p$ may be either 'i' later or 'ii' earlier (including equal to) than $T_{eg} + E_n$. In case 'i', epoch length will be regenerated randomly at $T_{eg} + E_n$. However, in case 'ii', the memoryless property loses completely since the epoch length in this case is no longer random relative to T_p. This factor affects the accuracy of the above link availability estimation based on the memoryless property which ensures that the epoch length from t_0 onwards will still follow distribution $E(x)$ with the same λ^{-1}.

It seems likely that the above loss of the memoryless property can be remedied by regenerating randomly the epoch lengths for the related nodes at t_0. However, doing so violates the assumption on the independence of node movements since it makes the related nodes to change their epochs at the same time. Here, we are not to discuss a fully amended formula for Eq. 9 subject to the above mentioned factor since it will not happen in a real environment with exponential distributed epoch lengths. Nevertheless, to verify the above analysis, a partial amendment to Eq. 9 is still discussed below.



En = length of the next epoch, Tp = predicted link-available period

Fig. 2. Simulation model for epoch generation and T_p prediction

It is much easier to amend the formula for $L_1(T_p)$ than that for $L_2(T_p)$ due to the same reason as for calculating $L_2(T_p)$ mentioned earlier. Here, we only discuss an amended $L_1(T_p)$ subject to the loss of the memoryless property. As mentioned earlier, $L_1(T_p)$ is the probability for the epoch lengths to keep unchanged between t_0 and $t_0 + T_p$. Due to the memoryless property, $L_1(T_p) = [1 - E(T_p)]^2$. When this property loses as illustrated in Fig. 2, this probability should be equal to $[1 - E(l + T_p)]^2$ since the epoch length is randomly generated at T_{eg} rather than t_0, where l is the offset between T_{eg} and t_0.

TABLE I

MEASURED ϵ AND l FOR MOBILITY MODELS WITH EXPONENTIALLY DISTRIBUTED EPOCH LENGTHS

Mean epoch	$\lambda^{-1} = 60s$		$\lambda^{-1} = 250s$	
Space (m^2)	500^2	1000^2	500^2	1000^2
Fig. no.	1	3	4	5
ϵ	0.2808	0.2351	0.3036	0.2520
$l(s)$	7	7	15	13
Fig. no.	6	7	8	9
ϵ	0.2832	0.2377	0.5549	0.5319
$l(s)$	9	7	20	18

So, an amended $L(T_p)$, $L_a(T_p)$, is approximated by

$$\begin{aligned} L_a(T_p) \quad &\approx \quad e^{-2\lambda(l+Tp)} + \frac{1}{2\lambda T_p} + \epsilon \\ &+e^{-2\lambda T_p}(p\lambda T_p - \frac{1}{2\lambda T_p} - \epsilon - 1). \qquad (15) \end{aligned}$$

The dashed line in Fig. 1 is plotted according to $L_a(T_p)$ given by Eq. 15 with an experimental $l = 7s$[4], showing a tendency similar to the simulation results. Since the above mentioned case 'ii' happens relatively frequently in the region of small T_p (compared to λ^{-1}), this mismatch mainly appears for small T_p region. To reduce the inaccuracy to the measured ϵ caused by this mismatch, we avoid to measure ϵ for very short T_p. Our experience is to measure ϵ from $T_p \geq 0.5\lambda^{-1}$. Table I lists some measured ϵ and l for mobility models with exponentially distributed epoch lengths used in the paper.

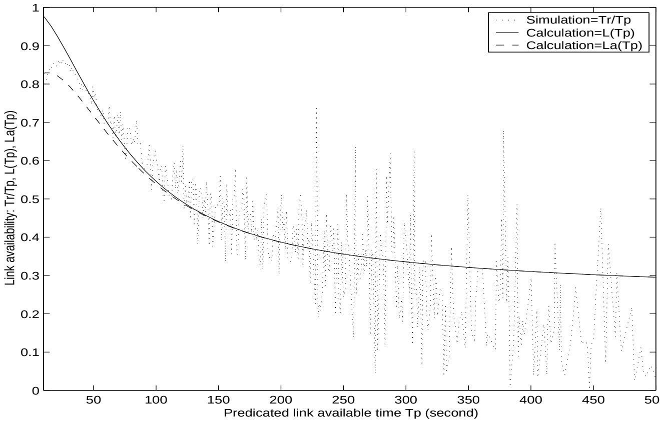


Fig. 3. Random walk-based model: $\lambda^{-1} = 60s$ and $(1000m)^2$

Fig. 3 shows some results for a space of $1000 \times 1000m^2$, which give a lower node density than $500 \times 500m^2$. It can be found that the calculation results are a little higher than those of simulation (on average) when T_p is large. This is because the number of events for large T_p values is much less than that for small T_p values with short λ^{-1} as shown in the figure. For example, $P\{$Epoch length > 350 with $\lambda^{-1} = 60\} = 1 - E(350) = 0.0029$ while $P\{200 \geq$ Epoch length > 100 with $\lambda^{-1} = 60\} = E(200) - E(100) = 0.1532$. As mentioned above, the ϵ used in the calculation is measured over the simulation results. It is true for all measurements that the final measured result tends to reflect frequently occurring events. So, the above phenomenon almost disappears in the simulation with a long mean epoch length as shown in Figs. 4 $\sim$ 5, in which more events for large T_p values are generated with $\lambda^{-1} = 250s$. For example, $P\{$Epoch length > 700 with $\lambda^{-1} = 250\} = 0.061 >> 0.0029$. However, the calculation results given by $L(T_p)$ in Fig. 4 trends to go beyond 1 when T_p becomes small. It is because the final measured ϵ has not taken into account the results for small T_p values as mentioned earlier. When mean epoch length becomes longer, the case 'ii' mentioned above will happen more frequently and cause a larger mismatch between the simulation and the calculation given by $L(T_p)$. $L_a(T_p)$ almost amends this problem as shown in these figures. Despite the above case for small T_p values, we can find that, the calculation results can generally reflect the tendency of the simulation results.

From Table I, we can also find that ϵ decreases with space size increases. This indicates that the probability for an available link to be still available after a movement change becomes small as node density decreases. So, in the case of uncertainty about ϵ, it is fine to set $\epsilon = 0$, i.e., using Eq. 14 for link availability estimation.

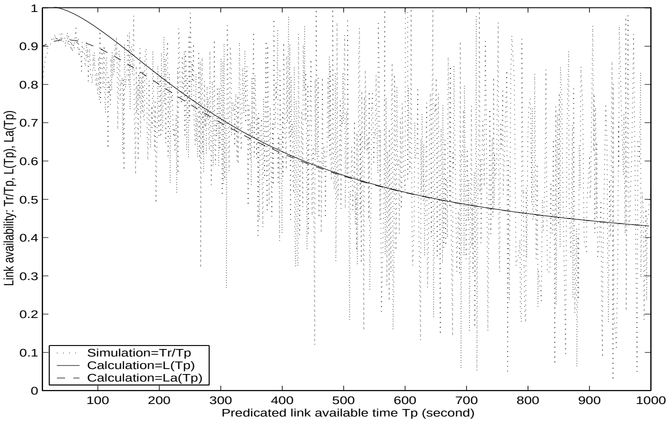


Fig. 4. Random walk-based model: $\lambda^{-1} = 250s$ and $(500m)^2$

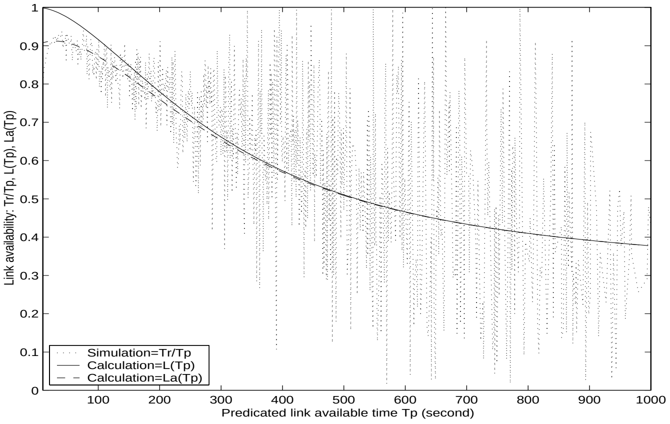


Fig. 5. Random walk-based model: $\lambda^{-1} = 250s$ and $(1000m)^2$

Now we look at some results for a modified random way point model with exponentially distributed epoch lengths (the epoch length distribution of the original model is non-exponential [13]).

[4] The experimental l is obtained by matching calculation results with simulation results. Since l will not be used in a real environment with exponentially distributed epoch lengths as mentioned above, here we do not discuss how to determine mathematically l only for simulation aspect.

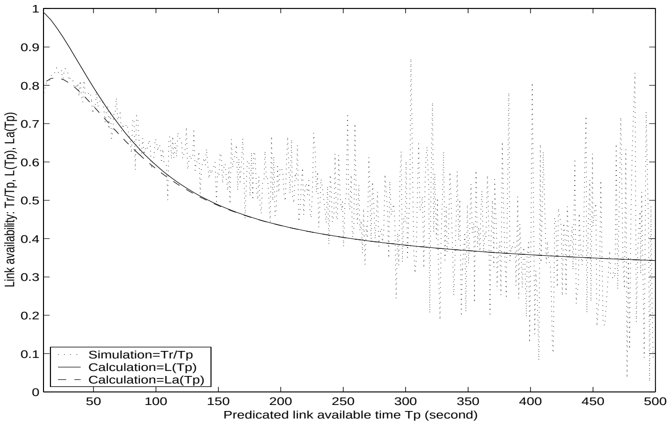


Fig. 6. Modified random way point model: $\lambda^{-1} = 60s$ and $(500m)^2$

In this model, a node selects uniformly a random destination point within a given space. The time for the node to reach the destination (i.e., epoch length) is selected according to an exponential distribution with mean λ^{-1}. Then the node's speed is the ratio of the distance to the destination over the epoch length. Upon reaching the destination, the node first stays there for a *pause* seconds (here $pause = 0$) then repeats the above operation. During the movement to the destination, the node's velocity keeps unchanged. This model has two differences from the random walk-based model: a node can only move within the given space and the speed is non-uniformly distributed without a predefined limitation.

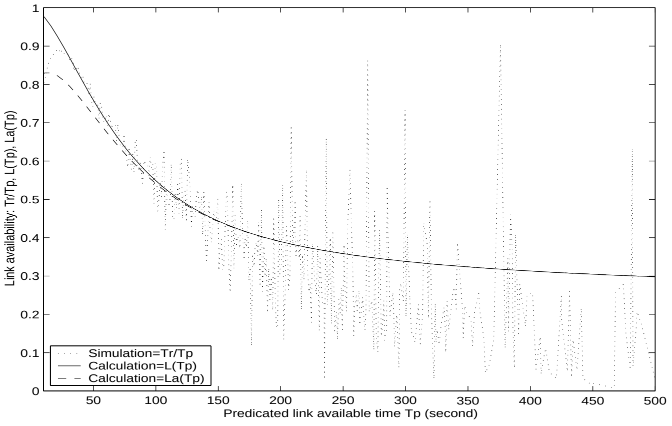


Fig. 7. Modified random way point model: $\lambda^{-1} = 60s$ and $(1000m)^2$

We first look at some results for $\lambda^{-1} = 60s$. As shown in Figs. 6 $\sim$ 7, the phenomena here are similar to those for the random-walk based model for respective space sizes, and so is for the ϵ values as shown in Table I. However, in the case of $\lambda^{-1} = 250s$, there is a big difference in ϵ between these two models. As mentioned earlier, a larger ϵ value indicates a larger probability for an active link between a pair of nodes to be continuously available after changes in movements. So, this difference means that this probability with the modified random way point model is larger than that with the random walk-based model. This is because the random walk-based model allows a node to move beyond the boundary of a given space, resulting in a node density lower than that with the modified random way point model. Furthermore, the time for a node to be outside of a given space is proportional to its epoch length which depends on λ^{-1}. This can explain why the above difference is not so obvious in the case of $\lambda^{-1} = 60s$.

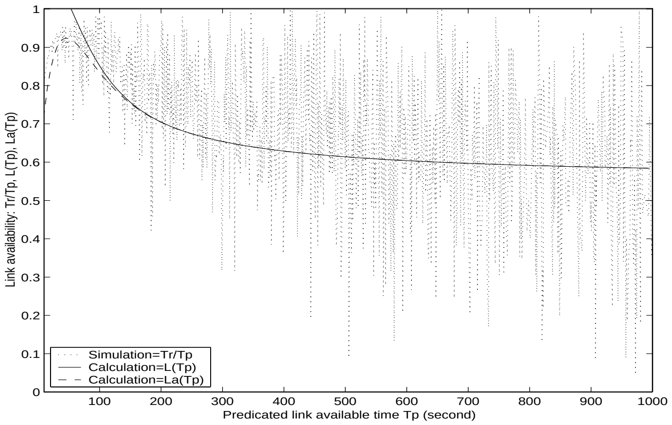


Fig. 8. Modified random way point model: $\lambda^{-1} = 250s$ and $(500m)^2$

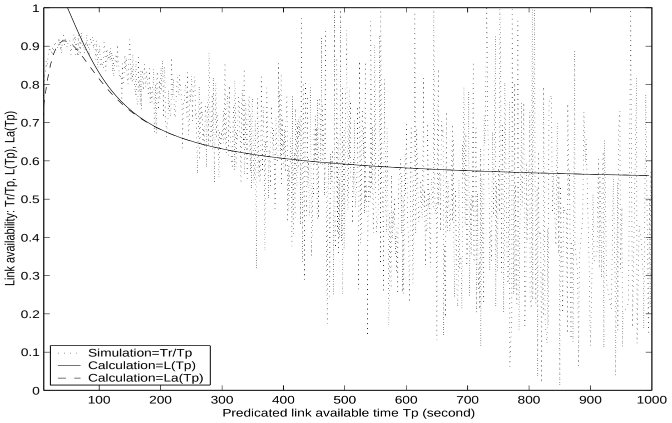


Fig. 9. Modified random way point model: $\lambda^{-1} = 250s$ and $(1000m)^2$

At last, we check the possible applicability of the proposed link availability estimation in environments with non-exponentially distributed epoch lengths. The mobility model adopted here is a so-called random way point model [13]. The only difference from the above modified random way point model is that here the speed is selected uniformly from 0 to $20\,m/s$. So, the epoch length is the ratio of the distance to the destination over the speed. In this case, it is difficult to quantify the epoch length distribution, and λ^{-1} cannot be initialized. We obtain λ^{-1} values for three space sizes through statistics on the simulation results.

It is not strange that there is a big mismatch between the calculation results and the simulation results in this case as shown in Figs. 10 $\sim$ 11. It is interesting that this mismatch tends to become smaller as space size increases, and the minimal link availability, $L_{min}(T_p)$ given by Eq. 14 (circled line in the figures), can still be suitable for this case (on average). Of course, the link availability estimation for non-exponentially distributed epoch lengths still needs more studies.

IV. A ROUTING METRIC BASED ON $L(T_p) \times T_p$

One practical interest of the proposed estimation is to develop a path selection metric in terms of path reliability because T_p alone cannot properly judge link availability. For example, if

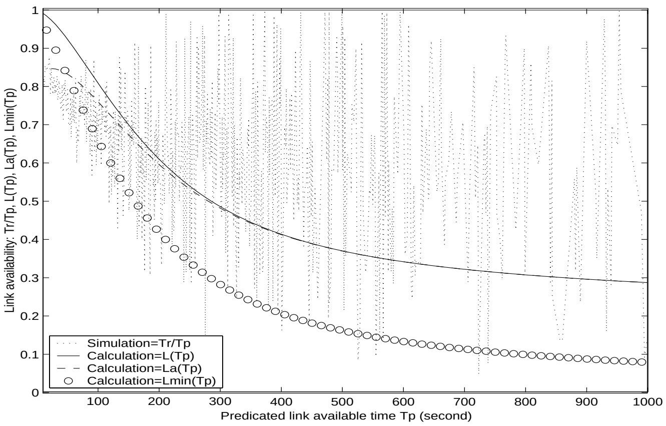


Fig. 10. Random way point model for space $500 \times 500m^2$ (non-exp, $\lambda^{-1} = 161s$, $\epsilon = 0.2084$ and $l = 15s$)

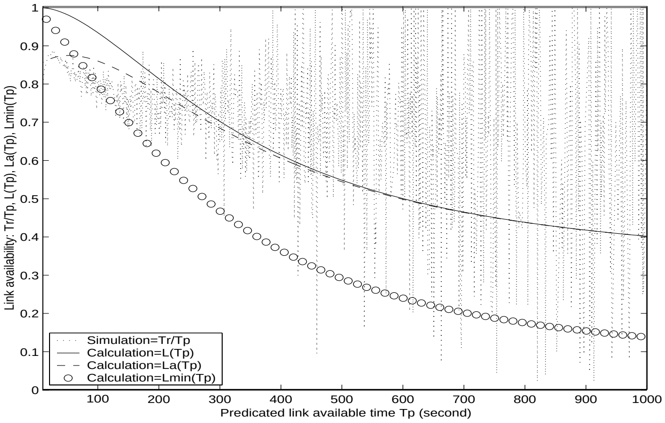


Fig. 11. Random way point model for space $1000 \times 1000m^2$ (non-exp, $\lambda^{-1} = 276s$, $\epsilon = 0.2639$ and $l = 23s$)

there are two links: $\mathbf{L}1$ with $T_p = 10$ and $L(T_p) = 0.001$ while $\mathbf{L}2$ with $T_p = 5$ and $L(T_p) = 0.1$. Only using T_p as the metric to select a more reliable link, $\mathbf{L}1$ will be selected. But its mean available time $10 \times 0.001 = 0.01$ is much smaller than $5 \times 0.1 = 0.5$ of $\mathbf{L}2$. Therefore, it is better to use $L(T_p) \times T_p$ as the metric.

For a path consisting of multiple links, its reliability mainly depends on the weakest link in terms of reliability. The link reliability can be judged in terms of $\overline{T_r} \approx L(T_p)T_p$. A link with smallest $\overline{T_r}$, i.e., $\overline{T_r}_{min}$, is considered as the weakest link. An ideal path in terms of path reliability should be the one with a maximal $\overline{T_r}_{min}$ as well as a minimal number of links. In the following, we show through computer simulation some performance differences given by the dynamic source routing (DSR) scheme [16] using the above routing metric and classical routing metrics such as ‘first found path’ and ‘shortest path’ which are defined below. The same simulation model as used in Section III is adopted here. A source node transmits data packets at fixed packet rates, and packet length is also fixed at 64 bytes. The following routing metrics are adopted in the simulation:

- *First found path (FFP)*: The first found path is selected. For example, in DSR, the path carried by the first routing reply packet received by the source node is selected.
- *Shortest path (SP)*: The path going through the minimal number of hops is selected.
- *Shortest path with link reliability estimation*: This is what we have defined above. This metric can be further classified according to link availability estimation algorithms as follows:

-*spLREo:* $\overline{T_r} \approx L(T_p)T_p$ using $L(T_p)$ given by Eq. 10;
-*spLREc:* $\overline{T_r} \approx L_{min}(T_p)T_p$ using $L_{min}(T_p)$ given by Eq. 14;
-*spLREa:* $\overline{T_r} \approx L_a(T_p)T_p$ using $L_a(T_p)$ given by Eq. 15.

TABLE III

COMPARISON FOR THE RANDOM WAY POINT MODELS: 5 FLOWS, $10\, packets/s/flow$, $(1000m)^2$

M-models	R-metric	loss	delay (s)	goodput	+GP(SP)
Modified:	*SP*	0.6009	1.379	19.96	NA
$\lambda^{-1} = 60s$	*spLREc*	0.5954	4.039	20.23	1.4%
Original:	*SP*	0.0601	0.030	46.99	NA
$\lambda^{-1} = 276s$	*spLREc*	0.0634	0.040	46.83	-0.3%
Original:	*SP*	0.0357	0.096	48.22	NA
$\lambda^{-1} = 56s$	*spLREc*	0.0296	0.162	48.50	0.6%

M-models: mobility models

The performances given by routing metrics are measured in the following terms:

- *User data packet loss ratio*: The number of lost packets over the number of transmitted packets;
- *Average end-to-end delay*: The average delay experienced by the received user data packets;
- *Goodput*: The number of received packets over the simulation time ($packets/s$).

Table II lists some simulation results for the random walk-based model. All the five routing metrics listed above have been investigated for two packet transmission rates. We can find that the performances given by *spLREx* (x=‘o’,‘c’ and ‘a’) are the best in terms of goodput. Since the delay is measured only for the received packets, it becomes shorter when packet loss ratio reaches some values. It is because DSR will drop the buffered packets that have been queued for certain threshold (here $30s$). When a path is down, rerouting operation may contribute significantly to the queuing delay, resulting in more long-delayed packets to be dropped, and this phenomenon becomes more obvious in the case of heavy load as shown in the table.

Column ‘+GP (FFP,SP)’ shows the goodput improvement given by *spLREx* over *FFP* (the left one) and *SP* (the right one), respectively. We can find that this improvement increases with packet load. However, there is almost no difference in the performances given by *spLREx* especially in the case of $10\, packets/s$. This is because, although the three metrics give different values of $\overline{T_r}$, they are almost similar in reflecting the degree of path reliability so that each metric may select the same reliable path. In the following comparison, we only compare *SP* with *spLREc* since it is simple to implement *spLREc* due to the simplicity of $L_{min}(T_p)$ calculation as mentioned in Section II.

Table III lists some results for the modified and original random way point models. We can find that the differences in goodput given by *SP* and *spLREc* are very small. It is because in these cases, mobility is almost not the major concern for routing in terms of path reliability because the random way point model has a higher connectivity than the random walk-based model since the latter allows a node to move beyond the boundary of a given space as mentioned in Section III. Therefore, the

TABLE II

COMPARISON FOR THE RANDOM WALK-BASED MODEL: 5 FLOWS, $\lambda^{-1} = 60$, $(500m)^2$

Packet rate	10 packets/s/flow				20 packets/s/flow			
R-metric	loss	delay (s)	goodput	+GP (FFP,SP)	loss	delay (s)	goodput	+GP(FFP,SP)
FFP	0.6613	45.03	16.94	NA	0.8562	3.510	14.38	NA
SP	0.7104	24.29	14.48	NA	0.8700	19.99	13.00	NA
spLREo	0.6346	13.79	18.27	7.9%, 26.2%	0.6626	27.19	33.74	134.6%, 159.5%
spLREc	0.6346	13.79	18.27	7.9%, 26.2%	0.6564	30.21	34.36	138.9%, 164.3%
spLREa	0.6346	13.79	18.27	7.9%, 26.2%	0.6564	30.21	34.36	138.9%, 164.3%

probability for a path to break due to mobility is low in this case. $\lambda^{-1} = 276s$ in the original random way point model indicates that the frequency for a node to change its movement is much low, which further dilutes the effect of mobility on the performance. However, we can also find that there is still an improvement albeit small when $\lambda^{-1} = 56s$ although this model does not hold an exponentially distributed epoch length required by the proposed algorithm since L_{min} can somewhat reflect the link reliability of this model as mentioned in Section II.

For a given environment, *spLREc* can be further simplified into a metric just judging T_p instead of $\overline{T_r} = L(T_p) \times T_p$ to avoid calculating $L(T_p)$ every time. We can get $\frac{\partial \overline{T_r}}{\partial T_p}$ as

$$\frac{\partial \overline{T_r}}{\partial T_p} = e^{-2\lambda T_p}(1 + \lambda T_p - \lambda^2 {T_p}^2). \tag{16}$$

It is easy to get a T_p value for the peak of $\overline{T_r}$ against T_p, $T_{p,peak}$, as

$$\begin{aligned} T_{p,peak} &= \frac{1+\sqrt{5}}{2}\lambda^{-1} \\ &\approx 1.618\lambda^{-1}. \end{aligned} \tag{17}$$

It is also easy to show that $\overline{T_r}$ increases with T_p when $T_p \leq T_{p,peak}$ while decreases in the other cases. Therefore, an ideal link in terms of reliability should be the one with a maximal T_p in the case of $T_p \leq T_{p,peak}$ while with a minimal T_p in the other cases. However, when there are some T_p values smaller than $T_{p,peak}$ while some others larger than $T_{p,peak}$, $\overline{T_r}$ is still required.

Note that here we do not try to claim that *spLREx* is one best routing metric. Instead, what has been shown here is that the network performance (e.g., goodput) can be further improved with a proper consideration of path reliability in routing when mobility is the major concern for routing in terms of path reliability. By 'proper', we mean that some method to judge path reliability is necessary. The proposed link availability estimation and the corresponding routing metric are just such an effort. We have investigated this effort by combining the link availability estimation with the popular *SP* metric. Since, in general, the reliability of a path is reversely proportional to the number of hops or links it goes through, *SP* is better than other routing metrics in terms of path reliability. The performance improvement shown here indicates that the proposed link availability estimation can help *SP* in selecting a more reliable path.

V. CONCLUSION

We have proposed and investigated a prediction-based link availability estimation, $L(T_p)$, for MANET in this paper. This algorithm tries to predict the probability that an active link between two nodes will be continuously available for a predicted period, T_p, which is obtained based on the current node's movement. Although this algorithm cannot accurately calculate the link availability, it can reflect the general tendency of a link availability as shown by the simulation results, that is, $L(T_p)$ can approximate $\frac{\overline{T_r}}{T_p}$, where $\overline{T_r}$ is the mean time that a link will be continuously available corresponding to a prediction T_p. One practical interest of this algorithm is to develop a path selection metric in terms of path reliability, i.e., $L(T_p) \times T_p$, which can improve the network performance as shown by the simulation results. However, the link availability estimation for non-exponentially distributed epoch lengths needs further studies.

REFERENCES

[1] E.M. Royer and C.K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks", *IEEE Personal Communications*, (April. 1999), pp. 46-55.

[2] C.K. Toh, "Associativity-based routing for ad hoc networks", *Wireless Personal Communications*, Mar. 1997, pp.103-139.

[3] C.K. Toh, "Wireless ATM & ad-hoc networks", Kluwer, Nov. 1996.

[4] R. Dube, C. Raia, K-Y Wang and S. Tripathi, "Signal stability based adaptive routing (SSA) for ad hoc networks", *IEEE Personal Communications*, Feb. 1997, pp. 36-45.

[5] A. Bruce McDonald and T.F. Znabi, "A path availability model for wireless ad hoc networks", *Proc. IEEE WCNC*, New Orlean, USA, (Sept. 1999), pp.35-40.

[6] A. Bruce McDonald and T.F. Znabi, "A mobility-based framework for adaptive clustering in wireless ad hoc networks", *IEEE JSAC*, 17(8) (Aug. 1999), pp.1466-1487.

[7] K. Thompson, G. J. Mille and R. Wilder, "Wide-area Internet traffic patterns and characteristics", *IEEE Network* (Nov./Dec. 1997), pp.10-23.

[8] D.J. He, S.M. Jiang and J.Q. Rao, "Prediction of link availability in wireless ad-hoc networks", *The 1999 Singapore Pervasive Computing Conference*, The National University of Singapore, Singapore, (Nov. 17-19, 1999).

[9] D.J. He, S.M. Jiang and J.Q. Rao, "Link availability prediction model for wireless ad hoc networks", *Proc. 2000 International Conference on Distributed Computing System Workshop*, Taipei, Taiwan, (April 10-13, 2000), pp.D7-D11.

[10] W. Su and M. Gerla, "IPv6 flow handoff in ad hoc wireless networks using mobility predication", *Proc. IEEE GLOBOCOM*, Rio De Janeiro, Brazil (Dec. 5-9, 1999), pp.271-275.

[11] L. Kleinrock, "Queueing systems, volume I: theory", *John Wiley & Sons*, (1975) ISBN 0-471-49110-1.

[12] "Mathematical handbook", *People Education Publication House of China*, CBN 13012 0165, (1979), pp273.

[13] D.B. Johnson and D.A. Maltz, "Dynamic source routing in ad hoc wireless networks", *Mobile Computing*, Boston, USA (1996), edited by T. Imielinski and H. Korth, Kluwer Academic Publishers, pp.153-181.

[14] "NOKIA A020 wireless LAN access point", *http://www.nokia.com/corporate/wlan/point_a020.html*.

[15] MIL 3 Inc., "OPNET modeler/radio 5.1".

[16] J. Broch, D.B. Johnson and D.A. Maltz, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks", *IETF Internet Draft, draft-ietf-manet-aodv-03.txt* (June 1999).