# On the Predictability of Link Lifetimes in Urban MANETs

Stephan Bohacek     Andjela Ilic     Vinay Sridhara

bohacek@udel.edu    andjela@udel.edu    vsridhar@udel.edu

## Abstract

*Link lifetime prediction is occurs frequently in MANET routing protocols. For example, in assigning cache timeout values and route durations, routing protocols make implicit predictions of link lifetime. Some protocols even make explicit predictions. In this paper the predictability of residual link lifetimes is examined. This investigation is performed via simulations that make use of detailed models of urban areas. Propagation is modeled using beam-tracing.*

## 1 Introduction

Throughout all MANET routing protocols, prediction of link lifetimes is pervasive. For example, on-demand protocols often cache routes and links. The selection of a time for which these links and routes remain cached is an implicit prediction of the route or link lifetime. In the case of proactive protocols, neighbors are probed periodically. The time between probes is typically selected by the protocol designer and is related to the designers beliefs about link lifetimes. Several routing protocols make explicit estimates of link lifetimes. For example, in [1], an association is made between signal strength and longevity of the link. Specifically, a link is considered strong and suitable for use if an exponential average of the signal strength is above a threshold for a certain amount of time. In [2], the longevity of a route is estimated by observing the change in signal strength. In [3], a link is declared to be stable if it has lasted beyond a threshold where the threshold depends on the transmission range and the relative speed of the nodes. In [4], a method is presented that tries to predict the link time-to-live (TTL), i.e., the time until the signal strength will fall below a threshold by examining the signal strength and the rate of change in the signal strength. In [5], GPS is used to estimate the speed and time until the nodes move out of "range." In [6], the time that a link will last is based on GPS. Based on this prediction, the routing attempts to search for an alternative path before the link breaks.

While the idea of link lifetime and lifetime prediction is common, there has been little investigation of whether and the degree to which link lifetimes are predictable. Indeed, many of the references cited above did not attempt to justify their assumptions about link lifetime. Investigations that have focused on link lifetimes include [7], [8], and [9]. In [7] extensive simulations were performed. This work indicated that age (i.e., the time that the link has been alive so far) is a useful indicator of the residual link lifetime. In [8] and [9], the distribution of link lifetimes and the impact that link lifetime has on routing protocols were examined. It was found that link lifetimes and the way that protocols attempt to predict the lifetime plays an important role in the overall performance of routing protocols.

However, this previous research has not focused on the predictability of link lifetimes. Furthermore, all previous studies have assumed trivial mobility models and either free-space propagation or a two-ray model. While such models may have some applicability in desert or rural areas, there is considerable interest in using MANETs in urban areas. This paper examines the predictability of link lifetime in pedestrian MANETs in urban areas. Specifically, we study predictability of the link lifetime when the prediction is based on age, position, signal strength, relative speed of the end-nodes as well as many combination of these factors. Propagation is modeled using beam tracing. Several conclusions are made. In general, the predictability of link lifetimes is moderate; the best that can be achieved is a 60% reduction in the mean square error (MSE) of the prediction error when all the factors are utilized. In many cases, the difference between predictors that use the same number of factors is small. Furthermore, there are strong spatial and other heterogeneous effects. Thus, a predictor that works only marginally well in general, will, in some situations, work extremely well and produce dramatic reduction of the error variance. The differences between nodes inside of buildings and outside of buildings is significant with link lifetimes of nodes inside of buildings less predictable than links of nodes that are outside. Nodes in an free-space environment (two-ray propagation and random way-point mobility) have significantly more predictable residual link lifetimes. Several other differences between free-space and urban environments are also found.

The paper proceeds as follows. In Section 2, the simula-

**Figure 1. Cities used for Simulation. The rectangles are office buildings with evenly distributed offices and hallways (not shown). Outside sidewalks are shown. The mobile nodes are restricted to sidewalks, offices and hallways.**

tions used to reach these conclusions is discussed. In Section 5, several predictors that make use of a single factor are investigated in detail. In section 6, the performance of predictors that use a combination of several factors is presented.

## 2 Simulations

The conclusions reached in this paper are based on series of simulations. Four different environments were considered. One environment is a well-known random waypoint/free-space environment. We will refer to this environment as *free-space*. In this environment, the nodes picked random locations in a 1km×1km space and traveled to those locations at random speeds. In this investigation we focus on pedestrian networks. Thus the speed of the nodes were selected to match typical walking speeds, i.e., walking speed are Gaussian distributed with mean 1.34 m/s and standard deviation 0.26 [10]. The propagation in this first environment was the two-ray model

$$P_R = \begin{cases} C/d^2 \text{ for } d < 200 \\ C \times 200^2/d^4 \text{ for } d \geq 200 \end{cases}.$$

The transition at 200 m is consistent with a 1.5 m antenna height.

Three specific urban areas were examined, the University of Delaware Campus (UD), the Paddington area of London (Pad), and a hypothetical grid-city (GC). The first two cities are shown in 1. The grid city is a 1km×1km grid with each block 100m×85m. Roads and sidewalks separate the buildings by 30m. Each block has a single building.

In these cities, the simulations use realistic urban propagation and mobility models. Details of this model can be found in [11]. Briefly, the propagation is based on a 3-D ray-tracing technique. This model utilizes 3-D beam-tracing that include reflections off of the ground and of off buildings, transmissions into buildings, diffraction using the uniform theory of diffraction, and propagation through buildings with the attenuation factor model.. The transmission error probability is modeled to match typical 802.11b at 1Mbps with no reception possible if the channel gain is greater than -56dB. Thus, nodes are defined to be connected if the path loss between them is less than 56dB. A path loss of 56dB allows for free-space propagation of around 650 m.

The urban mobility model assumes that nodes (pedestrians) move from their home office, through the building, and, if the destination is in another building, the nodes move along sidewalks to the building, and then to the desire office destination. While walking, nodes may encounter congestion. In this case, the walking speed will vary as described in [11]. Also, if walking outside, nodes may come to traffic lights and be forced to wait. The traffic lights followed a 90 sec. period. If the node is not at a traffic light and not encountering traffic it maintains randomly selected speed that is constant for the duration of its trip. When a node reaches it the destination, the node waits for exponentially distributed amount of time before returning to the home office. Upon returning to the home office, the node waits an amount of time that is approximately exponentially distributed. In both cases, the mean waiting time is 1200 sec. (20 minutes). The nodes pick in-building destinations with probability 5/6 and out-of-building destinations with probability 1/6. The destinations selected are locations of meetings. Other nodes will also attend the meetings. Nodes tend to arrive when the meeting starts and leave when the meeting ends.

Each simulation was run for 50 hours of simulation time with 64 nodes. Approximately 3 million links were observed. Consequently the statistics computed are accurate. In most cases, the 95% confidence interval gave an error of well under 0.01%. However, it will be noticed that not all curves are smooth. We speculate that this is due to the heterogeneous structure of the environment. However, this question requires further study.

## 3 Methodology

We examine the ability to estimate the residual lifetime of a link, which we denote as $R$. The residual lifetime is the remaining lifetime of the link. The total lifetime of a link is $R + Age$ where $Age$ is the current age, i.e., how long the link has been alive so far. We say that a link life begins when the path loss between two nodes is less than or equal to 56dB and it ends when the path loss exceeds 56dB. In

this investigation, time is discretized at 1 sec. intervals. If the path loss had been below 56dB and then exceeds 56dB at the next time step, its residual lifetime is zero. And, if the path loss had been above 56dB and has just dropped below 56dB, then its age is 1 sec.

We seek to determine if various factors can be used to predict the residual lifetime. These factors include the age of the link, the relative speed to the nodes, the path loss over the link, and the location of the node. We use the relative mean square error (RMSE) to quantify the utility of the factor or factors. Specifically, letting $F$ represent a factor (or factors) used to predict $R$, we quantify the utility of this prediction via

$$\frac{E\left((R - E\left(R|F\right))^2\right)}{E\left((R - E\left(R\right))^2\right)}. \tag{1}$$

The choice of mean square error as the metric of interest is due to the fact that such a metric is widely used to quantify and compare predictors. However, the usefulness of this metric within the context of routing protocols is an area that is left for future research.

It is useful to examine the performance of a predictor when it is known that the link is to a node that is either inside or outside. To make such an examination we use the follow RMSE

$$\frac{E\left((R - E\left(R|F, \text{node is inside}\right))^2 \,\middle|\, \text{node is inside}\right)}{E\left((R - E\left(R|\text{node is inside}\right))^2 \,\middle|\, \text{node is inside}\right)} \tag{2}$$

$$\frac{E\left((R - E\left(R|F, \text{node is outside}\right))^2 \,\middle|\, \text{node is outside}\right)}{E\left((R - E\left(R|\text{node is inside}\right))^2 \,\middle|\, \text{node is outside}\right)} \tag{3}$$

Note that (1) implies that the predictor used is $E\left(R|F\right)$, which the optimal predictor when the metric is the mean square error. Also, the denominator of (1) is the variance, which can be thought of as a measure of the prediction error if the mean is used as a predictor. Thus, the RMSE is a comparison between a complicated predictor that uses the factor(s) to make the prediction and a simple predictor. Specifically, the RMSE is the fractional decrease in the prediction error variance of a complicated predictor compared to a error variance that results from using a simple predictor.

## 4 The Residual Lifetime of Links

The focus of this paper is on the residual lifetime of links. Figure 2 shows complementary cumulative distribution function (CCDF) of the residual lifetime of a link. The
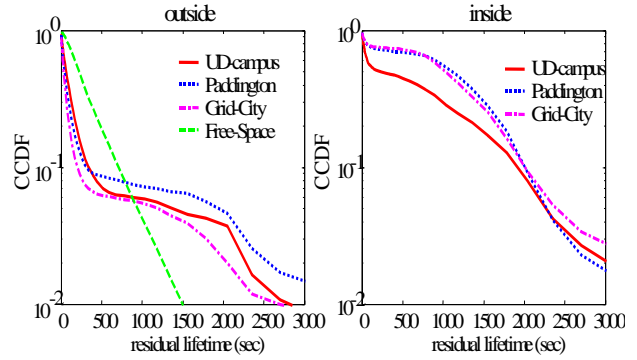


**Figure 2. CCDF of the Residual Lifetime of Links**

right-hand plot shows the distribution when the mobile node is inside a building, while the left-hand plot shows the distribution for nodes that are outside. The distribution resulting from each city investigated is shown along with the distribution from the free-space environment. Qualitatively, the distribution in the cities are similar to one another, whereas, the distribution from the free-space is quite different from any city. Furthermore, there is a significant difference between links to nodes that are inside and links to nodes that are outside. Specifically, nodes inside take part in links that have much longer residual lifetimes. The reason for this is that when a node is outside, it is moving (unless the node is at a traffic light or caught in sidewalk/hallway/doorway congestion). And, unless nodes are traveling to the same location, nodes will soon move far enough apart and/or entering a building so that communication between them is not possible. The duration of outside trips is noticeable by the steep decline of the CCDF for the residual lifetime less than 250 sec.

Nodes that are inside may be stationary for long periods of time. Thus, links to nodes that are inside may have long lifetimes. There are three subtle points about the nodes inside. First, as a node moves inside, it will walk down the hallway. As it passes near offices, it will be in communication range of the node sitting in the office. However, as it continues to move, this new link will be broken. These short lived links are visible in Figure 2 where between 20 and 40% of the links have a lifetime that is less than 40 sec. The reason that such a significant number of links have such a short lifetime is that a node has the possibility to communicate with a large number of nodes while moving down the hall, whereas when a node it is stationary, it will only make a few links to it nearby neighbors. A second subtle point is that nodes in meetings will be able to communicate with other nodes in the meeting. Since these nodes arrive at the same time, the duration of the connection will have

the same distribution as the duration of the meeting. A third point is that a node that is merely in its office will be able to communicate with nodes in nearby offices. However, such links may be broken when either of the nodes leaves their office. Thus, for these links, the mean link duration is half of the waiting time distribution.

Table 1 shows the mean and standard deviation for the residual lifetimes of links. The means shown are exactly the simple predictor referred to in Section 3. For each city, there are three types of links examined, namely, links from nodes that are inside, links from nodes that are outside, and links that are from nodes that could be outside or inside. For brevity, we often call links to node that are inside of buildings simply *inside links*. To gain a better understanding of how link behave, we focus on these three types of link throughout the paper. Thus, to examine the behavior of inside links, the "simple" predictor is the mean given in the top row of Table 1 and the denominator of (1) is the square of the second row from Table 1. We must stress that this division of links into these class is to assist in understanding the behavior of links. It is not a trivial task for a node to determine whether it is inside or outside. A truly simple predictor is one that does not include information as to whether the node is inside or not. The lower rows of Table 1 provide the mean and standard deviation of residual lifetimes where the node may be inside or outside. However, even these values assume that the city where the node resides is known. Predictors that are independent of the city is left for future research.

Table 1 shows the mean as well as the standard deviation. Notice that for links to nodes outside, the standard deviation is much larger than the mean. Thus, the coefficient of variation is far larger than one indicating high volatility. This volatility can also be observed in Figure 2 where the decay of the CCDF is seen to be slower than exponential. However, it is faster than polynomial and slower than the CCDF of a Weibull distribution. The reason for the high volatility is that nodes going to the same meeting may sometimes meet outside and then spend an extended period of time in the meeting.

The fourth column of Table 1 provides the mean and standard deviation for the free-space environment. Here we see that the coefficient is nearly one. Notice that the CCDF indicates that the residual lifetime is well modeled by an exponential distribution.

## 5 Predictability Based on a Single Factor

### 5.1 Age

From renewal theory, one of the first predictors that comes to mind is the prediction of the residual lifetime given the age, i.e., $E(R|Age)$. Figure 3 shows CCDF of

**Table 1. Statistics of the Residual Lifetimes of Links**

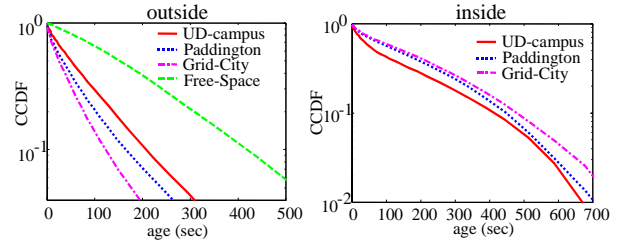| inside | UD | PAD | GC | OS |
|---|---|---|---|---|
| mean | 738 | 1131 | 1128 | - |
| std dev | 909 | 874 | 884 | - |
| outside | | | | |
| mean | 235 | 253 | 175 | 337 |
| std dev | 585 | 683 | 538 | 338 |
| both | | | | |
| mean | 487 | 692 | 652 | - |
| std dev | 805 | 898 | 873 | - |



**Figure 3. CCDF of Link Age**

the link age, while Figure 4 shows the expected residual lifetime given the current age. It can be seen that qualitatively, the expected residual lifetime conditioned on age is the same for the three cities. But inside and outside are distinct and free-space behaves different from any of the cities, both inside and outside. Specifically, for the free-space environment, the expected residual lifetime is nearly independent of the current age, but is slightly decreasing for links younger than 200 sec. On the other hand, in urban areas, the expected residual lifetime significantly increases with the age, hence older links tend to have a longer remaining lifetime than young links. The increase is significant; a link that is twice as old will have a residual lifetime that is between twice and four times greater. Thus, in urban areas, it is reasonable that a routing algorithm should select old links over young links. This contrasts the free-space environment where young links should be slightly favored over old links.

While the expected residual lifetime increases with age for both inside and outside links, the expected residual lifetime for outside links continuously increases, whereas for inside links, it rapidly increases and then stabilizes and slowly decreases. Thus, in the case of inside links, we see that links may experience "infant mortality," where new links have a strong tendency to break. As discussed above, these short lived links are from nodes moving through hallways. Nodes that are not in hallways but are stationary have links with longer residual lifetimes.
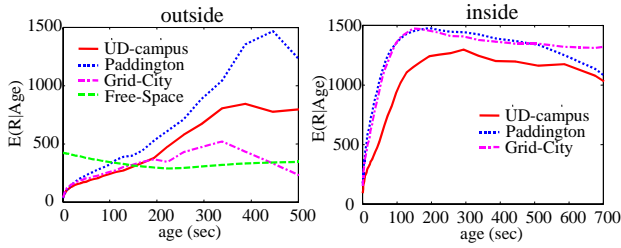
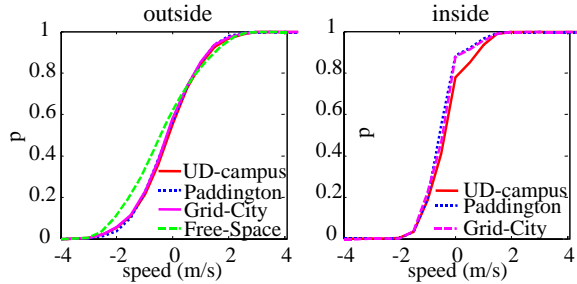**Figure 4. Expected Residual Lifetime given Age**



**Figure 5. CDF of Relative Speeds of Communicating Nodes**

Table 4 shows the performance of using $E(R|Age)$ to predict the residual lifetime of a link. (Section 5.5 discusses the meaning of the table.) We see that using age only provides a moderate decrease in the prediction error. For example, of links that are outside, conditioning on age results in decrease in the error variance by less than 10%. This contrasts Figure 4 that indicates that the residual lifetime is strongly dependent on the age. However, Figure 3 provide clarification and shows that there are a relatively small number of long lived links. Recall that this $E(R|Age)$ indicates a significantly different lifetime for old links as compared to young links. While this is true, most of the links that arise outside are young, hence the fact that old links will last longer is not particularly useful.

Table 4 does show that for links to nodes that are inside, predictions based on age are useful and can reduce the error variance by 20%. Similarly, links to nodes that are inside or outside can be predicted 15-25% better if age is used.

Table 4 also shows that in the free-space environment, conditioning on age has almost no impact. This is expected since $E(R|Age)$ is nearly constant.
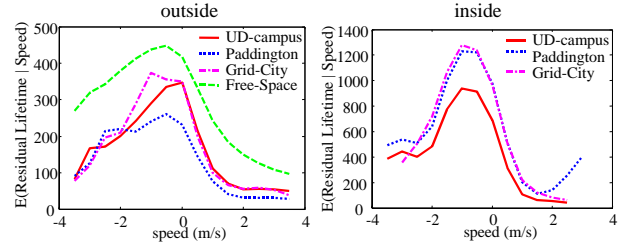


**Figure 6. Residual Lifetime and Relative Speed of Nodes.**

### 5.2 Speed

Next factor to be examined is relative speed between two nodes. Since only pedestrian mobility is considered, when nodes are moving their speeds are Gaussian with mean 1.36 m/s and standard deviation of 0.26. Thus, node speeds are typically below 2m/s. As a results, the relative speed between two nodes ranges from -4 m/s (if the two nodes are moving toward each other at top speed) to +4m/s (if the nodes are moving away from each other at top speed). Figure 5 shows the CDF of the relative speed. It can be seen that the distribution of relative speed is consistent across different cities and inside and outside.

Figure 6 shows the expected residual lifetime as a function of relative speed. In all environments, the predictors are qualitatively the same, the lifetime increases as the relative speed decreases for speeds that are greater than -1m/s. As the relative speed further decreases, the expected lifetime also decreases. It is intuitive that if nodes have a positive relative speed, the link between them will not last long. On the other hand, if nodes have a negative relative speed, then the nodes are moving toward each other. If the magnitude of the relative speed is large, then these nodes will soon reach locations where their distance is minimized and then begin to move apart resulting in the link breaking. Hence, as the relative speed decrease, the expected lifetime also decreases.

Table 4 shows RMSE if speed is used in prediction. The performance is fairly uniform across the different cities. For inside links and link that are either outside or inside, the MSE is reduced by 10-20%. For links that are outside, the decrease in MSE is minimal. Knowing the relative speed only decreases the error variance by 12% in the free-space environment.

### 5.3 Path Loss

If the path loss between two nodes is small (i..e, the signal strength is high), it seems unlikely that the link will
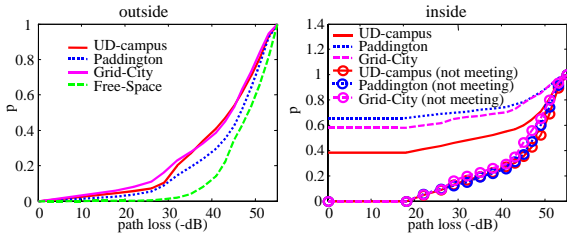
**Figure 7. Path Loss and Residual Link Lifetime. The axis is in -dB, hence the strongest signal is 0dB and weakest is -56dB (the right end of the x-axis). The right hand plot includes signal strength for all inside nodes as well as only for nodes that are not in meetings.**
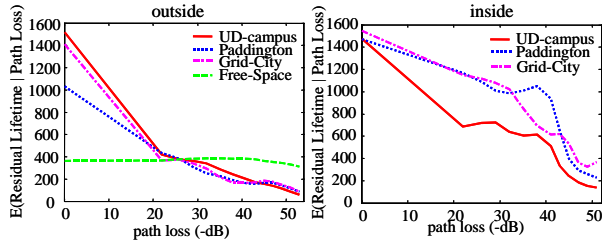


**Figure 8. Expected Residual Lifetime given the Path Loss.**

**Table 2. Dynamic Path Loss Based Prediction**

| inside | UD | Paddington | Grid | Free-Space |
|---|---|---|---|---|
| lag 0 | 0.57 | 0.66 | 0.68 | - |
| lag 0-1 | 0.56 | 0.65 | 0.68 | - |
| lag 0-2 | 0.56 | 0.64 | 0.67 | - |
| lag 0-3 | 0.55 | 0.62 | 0.66 | - |
| outside | | | | |
| lag 0 | 0.76 | 0.81 | 0.92 | 0.99 |
| lag 0-1 | 0.75 | 0.80 | 0.92 | 0.98 |
| lag 0-2 | 0.74 | 0.79 | 0.91 | 0.98 |
| lag 0-3 | 0.73 | 0.78 | 0.90 | 0.97 |
| both | | | | |
| lag 0 | 0.57 | 0.59 | 0.57 | - |
| lag 0-1 | 0.57 | 0.58 | 0.56 | - |
| lag 0-2 | 0.57 | 0.57 | 0.56 | - |
| lag 0-3 | 0.56 | 0.56 | 0.55 | - |

break in the near-term. However, it is unclear what a low path loss implies about the link lifetime in the more distant future. Furthermore, just because there is an intuition that a strong link will have a long life, it is not clear how often such strong links occur. In which case prediction based on path loss will perform poorly as did age-based prediction.

Figure 7 shows the CDF of the signal strength. For nodes that are outside, the signal strength is often quite weak. For example, 20% of the links have path loss that is within 6dB of where communication is not possible. This is intuitive since communicating with a distant node results in weak signal strength and there are more nodes that are far than nodes that are close. For nodes that are inside, we see two types of behavior caused by the fact that some nodes are in meetings and hence in close proximity of other nodes. Figure 7 shows that nodes that are not in meetings have a path loss distribution that resembles the path loss distribution of nodes that are outside. However, when all inside nodes are considered, a significant number of nodes have a very small path loss. Clearly, these are links between nodes that are attending the same meeting. As mentioned (and is well known), meetings can last a long time, and hence the links last a significant amount of time and thus can make up a large fraction the observed links.

The distributions of the path loss of links to nodes outside are consistent across all three cities and not significantly different from free-space. For nodes inside and not in a meeting, the distribution of path loss are nearly the same for the three cities. However, when meetings are considered, the distributions are quite different.

Figure 8 shows the expected residual lifetime given the path loss. The results agree with intuition; the stronger the signal, the longer the link will last. For links to nodes inside and outside, the expected residual lifetime reaches around 1400 sec., when the path loss is very small. For the inside nodes, these are nodes that are in a meeting together and can make up a significant portion of all inside links. For nodes that are outside, these links are between nodes that are going to the same meeting. However, as Figure 7 shows, this rarely occurs.

Figure 8 also shows the expected residual lifetime for links in the free-space environment. This figure indicates that the lifetime of a link is not strongly related to the strength of the signal.

Table 4 shows the RMSE when path loss is used in the prediction. We see that in many cases, the prediction error is greatly reduced with this predictor. Indeed, path loss is the best performing predictor. For nodes that are inside or outside, the error variance is reduced by over 40%. In most other cases, the reduction is also significant. However, the prediction of the residual lifetime for outside nodes in the grid city is only reduced by 8%. In the case of free-space, the path loss provides almost no reduction in the prediction

error.

As mentioned in Section 1, some have suggested using the change in the path loss or signal strength to predict the link lifetime. More generally, one can consider not just the current path loss, but the past path loss as well. Table 2 shows the results of using the current path loss and the past three path loss observations (path loss is sampled every second). Specifically, Table 2 shows the RMSE when only the current path loss observation is used, when the current observation of path loss is used along with the path loss observed one second ago, and so on. Thus, the table shows the impact of using more and more information. We see that while using the past path loss does decrease the error variance, the decrease is small.

## 5.4 Location

The performance of the predictors examined so far have been greatly influenced by whether the node is inside or outside. Here we will examine the impact of location in more detail.

Urban environments are spatially heterogeneous. While buildings cause a great reduction in the signal strength when signals pass through buildings, they also can act to guide the signals. For example, it is often the case that a street lined with tall, regular shaped buildings attenuates the signal less than it occurs in free-space propagation. Furthermore, when nodes are inside, they are often not moving, whereas when they are outside, they are moving (unless they are stopped at the traffic light). In this section we examine the impact that location of the node has on residual lifetime of the link. We consider information as to whether the node is inside or outside. This information is denoted by $In/Out$. We also consider more detailed information denoted $Loc$, which includes if the node is inside or outside and, if it is outside, it includes an approximate position of the node. The resolution of outdoor position was set to 20 meters.

Table 3 shows the RMSE for the location-based prediction (i.e., $E(R|In/Out)$ and $E(R|Loc)$). In comparison to the predictors discussed above, the performance of location-based prediction is poor. However, the error shown in Table 3 is averaged over the entire space. Thus, it is dominated by the behavior in locations that nodes often visit. Since the environment is heterogeneous, we can expect that location-based predictors will preform better in some locations than others. Figure 9 shows how the predictor $E(R|Loc)$ performs in different locations. Specifically, Figure 9 shows $\rho(Loc)$ where

$$\rho(Loc) := \frac{E\left((R - E(R|Loc))^2 |Loc\right)}{E\left((R - E(R))^2\right)}.$$

We see that in many locations the predictor does not de-



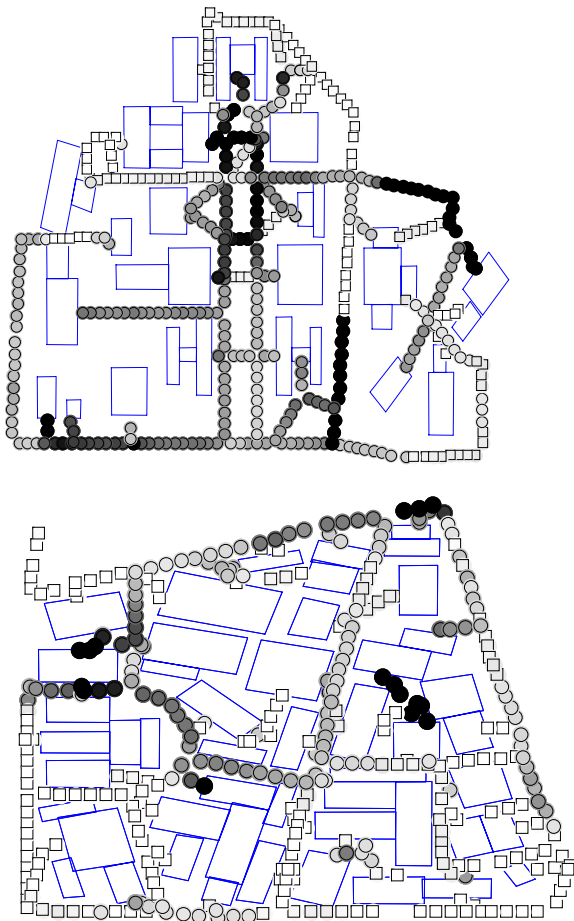**Figure 9. Performance of Location Based Predictor. Gray dots are placed along the sidewalk according to the reduction in variance conditioned on the location. The dark circles indicate no reduction in the variance. The light colors are greater reduction in the variance. The white squares indicate the variance is reduced by a factor of 10. The left hand plot is of the UD campus and the right hand plot is of Paddington.**

**Table 3. Location-Based Prediction**

| outside | UD | PAD | GC |
|---|---|---|---|
| $E_{out}(R\|Loc)$ | 0.97 | 0.90 | 0.92 |
| $E_{out}(R\|Age)$ | 0.91 | 0.89 | 0.97 |
| $E_{out}(R\|Age, Loc)$ | 0.68 | 0.59 | 0.74 |
| $E_{out}(R\|PL)$ | 0.74 | 0.80 | 0.91 |
| $E_{out}(R\|PL, Loc)$ | 0.62 | 0.56 | 0.69 |
| both | | | |
| $E(R)$ | 1 | 1 | 1 |
| $E(R\|Loc)$ | 0.89 | 0.78 | 0.73 |
| $E(R\|In/Out)$ | 0.90 | 0.80 | 0.74 |
| $E(R\|Age)$ | 0.83 | 0.77 | 0.74 |
| $E(R\|Age, Loc)$ | 0.67 | 0.60 | 0.74 |
| $E(R\|Age, In/Out)$ | 0.74 | 0.65 | 0.61 |
| $E(R\|PL)$ | 0.56 | 0.57 | 0.54 |
| $E(R\|PL, Loc)$ | 0.51 | 0.49 | 0.48 |
| $E(R\|PL, In/Out)$ | 0.55 | 0.53 | 0.53 |



**Figure 10. Expected Residual Lifetime given Age at Different Locations.**

crease the MSE (shown with the dark circles). However the lighter circles indicate areas where the predictor performs well. The squares indicate areas that the MSE is reduced by a factor of 10. Therefore, we can conclude that while the average user might not get substantial benefit from these location-based predictors, some users will.

While the location information does not, on its own, produce a good predictor, when joined with other factors, the predictors are much better. Table 3 shows that if age, which was found not to produce a good predictor, is combined with location, then a good predictor is formed. To understand this, consider Figure 10, which shows the expected residual age conditioned on age at different locations. The heterogeneity of the environment is obvious. The behavior near to the building entry is explained as follows. If a node is near to the building entry, either it will soon enter the building or has just left the building. In the first case, once it enters, all the links to other outside nodes will be broken. Moreover, if the node will soon enter the building, then its links are likely to be old since it had formed and maintained these links during its journey outside. On the other hand, a node that has just exited the building, has also just established links with other nodes that are outside. In this case its links will be quite young and will be maintained for quite some time as the node travels outside. The other curves in Figure 10 can be explained in a similar way.

### 5.5   Comparison of Single Factor Predictors

Table 4 shows the performance of the predictors that only use a single factor. The performance is examined in three ways, namely, the performance when it is known that the nodes are inside, the performance when it is known that the
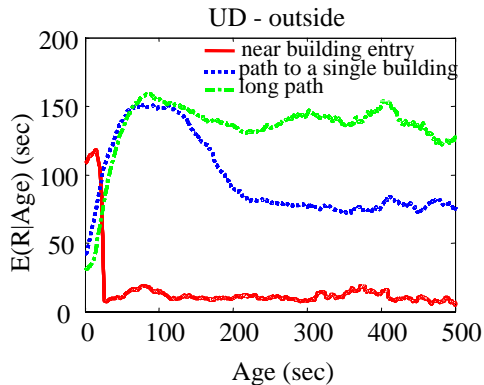
nodes are outside, and the performance when the nodes may be inside or outside (both). When all nodes are considered, the metric (1) is used, when only inside nodes are considered, metric (2) is used, and when only outside nodes are considered, metric (3) is used. The value of the RMSE is shown in the right three columns. Note that some predictors do not make sense in certain settings. For example, a predictor that uses the knowledge of whether nodes are inside or outside cannot be evaluated on inside nodes only.

The first column of Table 4 shows the predictor used. The subscript in or out denotes that predictor used the fact that all the nodes are inside or outside, respectively. The other columns provide the value of appropriate metric for the given predictor.

The performance of different predictors is discussed above. As mentioned, predictors based on path loss outperform other predictors in all settings except for free-space where none of the predictors perform particularly well.

## 6   Prediction using Multiple Factors

Section 5.4 showed how the combination of two prediction factors, age and location, can be combined and result in a predictor that performs far better than predictors based on the individual factors. Here we examine more predictors that use multiple factors.

Table 5 shows the value of RMSE for predictors that use three different factors. As expected, these predictors perform better then the predictors that use only a single factor. In many cases, we see that combining of two factors that, by themselves provided only a marginal reduction in the MSE, together provide more substantial reduction in MSE. It is for this reason that we include an examination of path loss along with the path loss observed one second ago. Predic-

**Table 4. Comparision of Single Factor Predictors**

| inside | UD | PAD | GC | OS |
|---|---|---|---|---|
| $E_{in}(R|Age)$ | 0.78 | 0.79 | 0.79 | - |
| $E_{in}(R|Speed)$ | 0.85 | 0.82 | 0.84 | - |
| $E_{in}(R|Path\ loss)$ | 0.57 | 0.66 | 0.68 | - |
| $E_{in}(R|Loc)$ | - | - | - | - |
| outside | | | | |
| $E_{out}(R|Age)$ | 0.90 | 0.89 | 0.97 | 0.98 |
| $E_{out}(R|Speed)$ | 0.96 | 0.96 | 0.97 | 0.88 |
| $E_{out}(R|Path\ Loss)$ | 0.76 | 0.81 | 0.92 | 0.99 |
| $E_{out}(R|Loc)$ | 0.97 | 0.90 | 0.92 | - |
| both | | | | |
| $E(R|Age)$ | 0.83 | 0.77 | 0.74 | - |
| $E(R|Speed)$ | 0.88 | 0.81 | 0.82 | - |
| $E(R|Path\ Loss)$ | 0.57 | 0.57 | 0.56 | - |
| $E(R|Loc)$ | 0.89 | 0.78 | 0.73 | - |

tors that use the past path loss are denoted with Path Loss+. Note that the inclusion of the past path loss had very small impact in Table 2, but when more factors are included, the impact is more significant. Nonetheless, considering past path loss observations only decreases the error variance a few percent.

When four factors are included, the MSE continues to decrease. Table 6 shows that path loss, along with detailed location information and age of the link provide the most reduction in the MSE. Table 7 shows progessively better predictors based on path loss. At each step, with more information added, the predictor substantially improves. Interestingly, age, which was shown to provide little predictive abilities by itself, substantially improves the performance of predictor when combined with path loss and location.

# 7 Conclusion

While link lifetime plays an important role in routing protocols design and performance, there has been little investigation into the predictability of link lifetimes. Here we examine many predictors in realistic urban environments. It is shown that link lifetime prediction mean square error can be reduced by nearly 60% as compared to using the mean lifetime as a prediction of the link lifetime. However, such predictors would require knowledge of the location of the node, the path loss across the link, and the age of the link. Simpler predictors are also useful. However, in all the cases, path loss leads to be the best predictor. The dynamics of path loss (e.g., the derivative of the path loss) provides little additional predictive abilities.

This investigation used the mean square error to quantify

**Table 5. Comparision of Predictors that use Three Factors**

| inside | UD | Pad | GR | FS |
|---|---|---|---|---|
| $E_{in}(R|Age,Speed)$ | 0.67 | 0.69 | 0.70 | |
| $E_{in}(R|Age,PthLos)$ | 0.52 | 0.56 | 0.62 | |
| $E_{in}(R|Speed,PthLos)$ | 0.51 | 0.57 | 0.62 | |
| $E_{in}(R|Age,PthLos+)$ | 0.49 | 0.54 | 0.60 | |
| $E_{in}(R|Speed,PthLos+)$ | 0.50 | 0.55 | 0.60 | |
| outside | | | | |
| $E_{out}(R|Age,Speed)$ | 0.80 | 0.79 | 0.91 | 0.86 |
| $E_{out}(R|Age,PthLos)$ | 0.71 | 0.76 | 0.90 | 0.93 |
| $E_{out}(R|Speed,PthLos)$ | 0.67 | 0.72 | 0.87 | 0.84 |
| $E_{out}(R|Age,PthLos+)$ | 0.67 | 0.71 | 0.87 | 0.93 |
| $E_{out}(R|Speed,PthLos+)$ | 0.64 | 0.68 | 0.85 | 0.93 |
| both | | | | |
| $E(R|Age,Speed)$ | 0.71 | 0.64 | 0.62 | - |
| $E(R|Age,PthLos)$ | 0.55 | 0.51 | 0.53 | |
| $E(R|Speed,PthLos)$ | 0.52 | 0.51 | 0.52 | |
| $E(R|Age,PthLos+)$ | 0.52 | 0.49 | 0.51 | |
| $E(R|Speed,PthLos+)$ | 0.52 | 0.49 | 0.50 | |
| $E(R|Age,I/O)$ | 0.74 | 0.65 | 0.61 | |
| $E(R|Speed,\ I/O)$ | 0.79 | 0.68 | 0.64 | |
| $E(R|Path\ Loss,\ I/O)$ | 0.55 | 0.53 | 0.52 | |
| $E(R|Path\ Loss+,\ I/O)$ | 0.56 | 0.50 | 0.49 | |

**Table 6. Comparision of Predictors that use Four Factors**

| inside | UD | Pad | GR | FS |
|---|---|---|---|---|
| $E_{in}(R|$Age, Speed,PthLos) | 0.46 | 0.51 | 0.58 | - |
| $E_{in}(R|$Age, Speed,PthLos+) | 0.44 | 0.48 | 0.56 | - |
| outside | | | | |
| $E_{out}(R|$Age, Speed,PthLos) | 0.62 | 0.64 | 0.81 | 0.81 |
| $E_{out}(R|$Age, Speed,PthLos+) | 0.58 | 0.57 | 0.76 | 0.81 |
| $E_{out}(R|$Age, Loc,PthLos) | 0.35 | 0.33 | 0.45 | - |
| both | | | | |
| $E(R|$Age, Speed,PthLos) | 0.49 | 0.46 | 0.48 | - |
| $E(R|$Age, Speed,PthLos+) | 0.47 | 0.43 | 0.47 | - |
| $E(R|$Age,PthLos,Loc) | 0.41 | 0.41 | 0.43 | - |
| $E(R|$Age,Speed,I/O) | 0.64 | 0.57 | 0.54 | - |
| $E(R|$Age,PthLos, IO) | 0.52 | 0.49 | 0.49 | - |
| $E_{out}(R|$I/O, Speed,PthLos) | 0.51 | 0.48 | 0.49 | - |
| $E(R|$Age,PthLos+, IO) | 0.50 | 0.46 | 0.48 | - |
| $E(R|$I/O, Speed,PthLos+) | 0.49 | 0.46 | 0.48 | - |
| $E(R|$Age,I/O Speed,PthLos+) | 0.44 | 0.41 | 0.44 | - |

**Table 7. Pregressively Better Path Loss Based Predictors**

| outside | UD | PAD | GC |
|---|---|---|---|
| $E(R|$PL) | 0.76 | 0.81 | 0.92 |
| $E(R|$PL,Loc) | 0.62 | 0.56 | 0.69 |
| $E(R|$PL,Loc,Age) | 0.35 | 0.33 | 0.45 |
| both | | | |
| $E(R|$PL) | 0.57 | 0.57 | 0.56 |
| $E(R|$PL,Loc) | 0.51 | 0.49 | 0.48 |
| $E(R|$PL,Loc,Age) | 0.41 | 0.41 | 0.43 |

the performance of a link lifetime predictor. However, the most relevant predictors should focus on network metrics such as overhead or throughput. Further work is required to understand the relationship between MSE and these network centric metrics.

# References

[1] R. Dube, C. D. Rais, K.-Y. Wang, and S. K. Tripathi, "Signal stability based adaptive routing (SSA) for ad-hoc mobile networks," *IEEE Personal Communication*, 1997.

[2] S. K. Das, A. Mukherjee, S. Bandyopadhyay, K. Paul, and D. Saha, "Improving quality-of-service in ad hoc wireless networks with adaptive multi-path routing," in *GLOBECOM*, 2000.

[3] C.-K. Toh, "Associativity based routing for ad hoc mobile networks," *wireless personal communication journal*, vol. 4, pp. 103–139, 1997.

[4] S. A. A. A. Ahila, P. Singh, and R. Shorey, "Route-lifetime assesment bases routing (RABR) protocol for mobile ad-hoc networks," in *Proc. IEEE International Conference on Communication 2000*, 2000.

[5] H. Dajing, J. Shengming, and R. Jianqiang, "A link availability prediction model for wireless ad hoc networks," in *Proceedings of the International Workshop on Wireless Networks and Mobile Computing*, 2000.

[6] W. Su, S. Lee, and M. Gerla, "Mobility prediction and routing in ad hoc wireless networks," *International Journal of Network Management*, 2000.

[7] M. Gerharz, C. de Waal, M. Frank, and P. Martini, "Link stability in mobile wireless ad hoc networks," in *Proc. Of the IEEE Conference on Local Computer Networks (LCN) 2002*, 2002.

[8] N. Sadagopan, B. K. F. Bai, and A. Helmy, "Paths: Analysis of path duration statistics and their impact on reactive manet routing protocols," in *MobiHoc*, 2003.

[9] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the impact of mobility on performance of RouTing protocols for ad-hoc NeTworks," in *Infocom*, 2003.

[10] G. K. Still, *Crowd Dynamics*. PhD thesis, university of warwick, 2000.

[11] V. Sridhara, J. Kim, and S. Bohacek, "Simulation of mobile ad hoc wireless networks," in *Submitted*, 2004.