CERIAS Tech Report 2007-49

Distance Reduction in Mobile Wireless Communication: Lower Bound Analysis and Practical Attainment

by Yu Dong, Wing-Kai Hon, David K. Y. Yau, and Jren-Chit Chin

Center for Education and Research in
Information Assurance and Security,
Purdue University, West Lafayette, IN 47907-2086

Distance Reduction in Mobile Wireless Communication: Lower Bound Analysis and Practical Attainment

Yu Dong, Wing-Kai Hon, David K. Y. Yau, and Jren-Chit Chin
Department of Computer Science, Purdue University, West Lafayette, IN 47907
{dong, wkhon, yau, jcchin}@cs.purdue.edu

***Abstract*— The transmission energy required for a wireless communication increases superlinearly with the communication distance. In a mobile wireless network, nodal movement can be exploited to greatly reduce the energy required by postponing communication until the sender moves close to a target receiver, subject to application deadline constraints. In this paper, we characterize the fundamental performance limit, namely the lower bound expected communication distance, achievable by any postponement algorithm within given deadline constraints. We consider a realistic map based stochastic movement model, of which the well known random waypoint model is a special case. For the random waypoint model, we develop a tight analytical lower bound of the achievable expected communication distance. In addition, we present a close analytical approximation of the lower bound that has a low computational complexity. For the general map based model, we characterize the lower bound distance experimentally. We also address the practical attainment of distance reduction (and hence, energy savings) through movement predicted communication. Specifically, whereas prior work has presented a least distance (LD) postponement algorithm and established its effectiveness experimentally, we provide an absolute performance measure of how closely LD can match the theoretical optimum. We show that LD achieves an average reduction in the expected communication distance within 62% to 94% of the optimal, over a realistic range of nodal speeds, for both the map based and random waypoint models. Moreover, the algorithm's absolute performance increases as the nodal speed or the allowable postponement delay increases.**

I. Introduction

To achieve energy efficient wireless communication, movement prediction [1] has been proposed to reduce the communication distance and hence communication energy for delay-tolerant applications. The basic idea is for a mobile sender to postpone communication, subject to given application deadlines, until a time when the sender is likely to move close to the receiver, or the communication target. Since in practice, the energy requirement of sending is proportional to the third or higher powers of the communication distance, the reduction in sending energy can be significant. In [1], several *postponement algorithms* are proposed to determine the best time of communication within application deadline constraints.

This paper is concerned with both the fundamental and practical performance of such energy-efficient movement-predicted communication. We consider a general and realistic *map-based* network model. In the model, a given geographical area is divided into a grid of fixed size cells. Nodal movement in the area can be regulated by given *accessibility constraints* modeling, for example, a map of freeways, roads, and streets (with possible speed limits) for terrestrial movement. Nodes move within accessible areas of the network in a succession of *trips*, each of which is defined by a starting and ending location. The exact route taken for each trip can then be specified by a given *route selection* algorithm. The route selection algorithm might similarly reflect how real people plan their road trips. For example, Internet tools like MapQuest and Yahoo Maps can return routes based on shortest travel time, most direct paths (say, major roads preferred with least number of road changes), etc. A model instantiation with null accessibility constraint and the selection of straight line paths in every trip would be similar to the well known random waypoint model [6], except that our grid based formulation will lead to a finite set of possible trip locations, whereas the original random waypoint model will have an infinite number of the possible locations.

For the random waypoint model, we derive tight lower bound expected communication distances achievable by *any* postponement algorithm, as a function of the average nodal speed and the allowable postponement delay. We also show how approximations of the lower bounds can be obtained by ignoring certain correlations in the sequence of cells visited within a trip. The approximation has a low computational complexity, but is remarkably close to the accurate bound in practice.

The lower bound results will allow us to fundamentally evaluate the performance of practical postponement algorithms. For example, several postponement algorithms are proposed and evaluated in [1]. Simulation and implementation experiments, for an enhanced version of the random waypoint model [9], show that a *least-distance* (LD) algorithm has the best practical performance despite its simplicity. Hence, whereas the prior work in [1] demonstrates the advantages of LD *relative* to competing algorithms, we for the first time provide an *absolute* performance measure of how closely LD can match the theoretical optimum. Our results show that LD achieves an average reduction in the expected communication distance within 62% to 94% of the optimal. Moreover, the algorithm's absolute performance increases as the nodal speed or the allowable delay increases.

Besides the analytical results, we present experiments to characterize the performance of movement prediction in a realistic instantiation of the map-based model to road travel in Lafayette, Indiana, USA. We also systematically evaluate how important system parameters such as the network grid size can affect algorithm performance.

A. Our contributions

The main contributions of this paper are as follows:

- We have developed a general map-based network and movement model to capture realistic nodal movement, while admitting the widely used random waypoint model as a special case. We have applied the map-based model to evaluate the performance of movement prediction in the Lafayette, Indiana area.
- We contribute to the understanding of fundamental performance limits in movement-predicted wireless communication. We have derived tight lower bound expected communication distances achievable by *any* postponement algorithm, as a function of the average nodal speed and the allowable postponement delay. Lower communication distances readily translate into higher energy savings for single hop wireless communication.
- We report extensive experimental results to verify and illustrate the analytical results. In particular, we quantify how closely the LD postponement algorithm in [1] can match the lower bound expected distances under different deployment scenarios. In addition, our experimental results systematically evaluate how important system parameters such as the network grid size can effect algorithm performance.

B. Paper organization

The balance of the paper is organized as follows. A general and realistic system model is given in Section II. Then we derive theoretical lower bounds of the expected communication distance in motion-predicted wireless communication in Section III. We review the basics of movement prediction and the LD algorithm in Section IV. Diverse experimental results verifying our theoretical analysis are presented in Section V. In Section V, we also quantify the ability of the LD algorithm in matching the upper bound distance savings, at various average nodal speeds. An approximation of the theoretical lower bounds of the expected communication distance is given in Section VI that has a low computational complexity. Related work is discussed in Section VII. Section VIII concludes.

II. Movement Model

In this section, we propose a general stochastic movement model that is based on the actual road maps to capture realistic nodal movements. Mathematically, the model is a quadruple $\langle \mathcal{N}, \mathcal{M}, \mathcal{T}, \mathcal{R} \rangle$ where $\mathcal{N}$ denotes the network configuration, $\mathcal{M}$ denotes the accessibility constraints similar to a map, $\mathcal{T}$ denotes the trip-based movement model, and $\mathcal{R}$ denotes the route selection algorithm. Details about each tuple of the model are given as follows.

Network configuration: In our model, a network is a two-dimensional X by Y rectangular area associated with a map of that area, where X and Y (in distance units) are the width and the height of the network, respectively. The whole network is divided into fixed size s by s square regions. Each square region is called a *cell*. Cells form a virtual grid over the network area, and each cell has a unique integer cell ID. To simplify boundary conditions, we assume that both X and Y are integer multiples of s. Thus the whole network has $m \times n$ cells, where $m = X/s$ and $n = Y/s$.

Accessibility constraints: The network is associated with a *map* defining the accessible areas of the network. In the map, a set of *pathways* (e.g., freeways, roads, and streets) may exist. These pathways constrain the routes between different locations in the network. Speed limits may be specified for each pathway.

Trip-based movement model: In our model, nodes move within the accessible areas of the network in a succession of *trips*, each of which is defined by a starting and an ending locations (i.e., the ending location of a trip forms the starting location of the next trip). The starting location of the first trip is chosen randomly from the whole area of the network under a uniform distribution. The ending location of a new trip is chosen upon reaching the ending location of the current trip, which is again chosen uniformly randomly from the whole area.

Once the end points of a trip are decided, they will be passed to a route selection algorithm (to be discussed below), which returns a sequence of pathways directing the mobile node from the starting location to the ending location. The actual speeds for the node to travel along each pathway are then chosen randomly between a minimum speed V_{min} and a maximum speed V_{max} (given as the speed limits of the pathway) under a uniform distribution.

Route selection algorithm: Given the starting and ending locations of a trip, the route selection algorithm generates a route between them that satisfies the accessibility constraints. Several map software and services (e.g., Google Maps, Microsoft MapPoint, Yahoo Maps, etc.) are available to provide map based route selection. Also, in case a chosen location happens to be in an inaccessible region, most route selection software automatically substitutes a nearest accessible location, and returns the corresponding route.

Thus, with a specific map of a given area (say a city or a state), our model may be used to generate movement patterns within the area.

Random waypoint model—A special case: If the accessible region is the whole network area, V_{min} in the trip-based movement model is set to 0, and the route selection algorithm always returns the direct line segment between the starting and ending locations, our map-based movement model is the same as the well known random waypoint movement model [6], except that our grid based formulation will lead to a finite set of possible trip locations, whereas the original random waypoint model will have an infinite number of the possible locations. By setting V_{min} to be greater than 0 instead, the random waypoint model is rectified so that it will reach a meaningful steady state average speed [9]. In this paper, apart from studying the general map-based movement model, we will also give a specific case study of the rectified random waypoint model.

To characterize the properties of the map-based stochastic movement model, we use the notations defined in Table I.

TABLE I
MAP-BASED STOCHASTIC MOVEMENT MODEL VARIABLES

Variable	Definition	Type
X	width of the network area	input parameter
Y	height of the network area	input parameter
s	cell size	input parameter
V_{max}	maximum nodal speed	input parameter
V_{min}	minimum nodal speed	input parameter
V	nodal speed	random variable
T	travel time	random variable
$E[V]$	expected speed	statistical property
$E[T]$	expected trip time	statistical property

III. Theoretical Lower Bound of Expected Communication Distance

In this section, we derive a method to obtain the theoretical lower bound of the expected communication distance under our movement model. We assume that there is a cell g in the network, which we call the *target cell* and contains a stationary receiver that we wish to communicate with. (We defer to Section V the experimental study of the lower bound when the receiver is mobile.) A mobile node moves around the network area from one cell to another during a trip, and the distance to the target cell may change accordingly. We want to calculate the expected distance between the target cell and the closest cell visited by the mobile in ℓ trips, which gives a *lower bound* on the expected communication distance between the mobile and the receiver during these ℓ trips.

To begin with, we recall that the whole network area is divided into a total of m by n cells, and each cell is a square of fixed size s by s. We now define $d(i, g)$, where $0 \leq i \leq mn - 1$, to be the Euclidean distance between a cell i and the target cell g; for simplicity, such a distance is measured between the centers of the two cells. Let N denote the number of distinct values of $d(i, g)$ over all i. We partition the cells into N sets so that cells in the same set have the same distance to the target. We denote each set of cells by S_j (where $0 \leq j \leq N - 1$), and we denote the distance between any cell in S_j and the target cell by D_{S_j}. For ease of discussion, we assume that the S_j's are sorted in increasing order of D_{S_j}; that is, $D_{S_j} < D_{S_{j+1}}$ for all j.

Now, consider the simple case where $\ell = 1$ (i.e., the mobile node takes only one trip). To find the expected shortest distance between the node and the target cell (which forms the lower bound on expected communication distance), we need to compute for each possible trip the different probabilities to attain a specific distance D_{S_j} as the shortest distance. After that, the desired expected distance can be computed in a straightforward manner.

To solve the case for a general ℓ, we can apply the same technique. To simplify our discussion, in the following subsection, we give a matrix representation that captures the probabilities of attaining the specific shortest distance D_{S_j} for all possible trips when $\ell = 1$, and based on that, we show how to calculate the expected shortest distance to the communication target in the general case.

A. Matrix representation of probabilities when $\ell = 1$

Assuming that the mobile node takes one trip, we define a two-dimensional matrix B to represent the probabilities of attaining a specific shortest distance for each pair of starting and ending locations as follows. The matrix B has $m \times n$ rows and $m \times n$ columns as illustrated below. Each element $b_{i,j}$ (where $0 \leq i, j \leq mn - 1$) represents a trip from cell i to cell j. Instead of storing a real-valued number as in the normal matrix definition, each element $b_{i,j}$ is a vector of size N. Each element $b_{i,j}[k]$, where $0 \leq k \leq N - 1$, defines the probability that the shortest distance is D_{S_k} for the trip from cell i to cell j. E.g., element $b_{2,3}[1]$ gives the probability that the shortest distance to the target is D_{S_1} for a trip from cell 2 to cell 3. Observe that for each element $b_{i,j}$ in B, the length-N vector has exactly one entry being 1, and all other entries being 0's.

$$B = \begin{pmatrix} b_{0,0} & \cdots & b_{0,j} & \cdots & b_{0,mn-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b_{i,0} & \ddots & b_{i,j} & \ddots & b_{i,mn-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b_{mn-1,0} & \cdots & b_{mn-1,j} & \cdots & b_{mn-1,mn-1} \end{pmatrix}$$

B. Matrix representation for a general ℓ

Assuming that the mobile node has traveled for ℓ trips, we want to compute the corresponding matrix representation, denoted by B^ℓ, of the probabilities of attaining a specific shortest distance for each possible pair of starting and ending locations of the travel. Then, in the next subsection, we show that the expected shortest distance can be computed easily based on this matrix representation.

Observe that attaining the shortest distance D_{S_k} after traveling ℓ trips with starting location i and ending location j occurs when the minimum of the two distances, namely (i) the shortest distance attained during the first $\ell - 1$ trips and (ii) the shortest distance attained at the last trip, is D_{S_k}. Let $E_{i,j,\ell}$ denote the event that the starting and ending locations of traveling ℓ trips is i and j, respectively, and let $D(i, j, \ell)$ denote the shortest distance attained by the corresponding travel. Then, based on the above observation, the probability $P(D(i, j, \ell) = D_{S_k})$—which is the probability of attaining shortest D_{S_k} after ℓ trips with starting location i and ending location j—can be expressed as:

$$\sum_{x=0}^{mn-1} P(E_{i,x,\ell-1}) P(\min\{D(i,x,\ell-1), D(x,j,1)\} = D_{S_k})$$
$$= \frac{1}{mn} \sum_{x=0}^{mn-1} P(\min\{D(i,x,\ell-1), D(x,j,1)\} = D_{S_k}),$$

where the last equality follows from the fact that $P(E_{i,x,\ell-1}) = 1/(mn)$, since destination of every trip is chosen uniformly randomly among all cells.

For the term $\sum_{x=0}^{mn-1} P(\min\{D(i,x,\ell-1), D(x,j,1)\} = D_{S_k})$, it can be computed if we have the matrix B and the matrix $B^{\ell-1}$, and the computation resembles a matrix multiplication. This suggests that the matrix B^ℓ can be defined recursively as follows:

$$B^\ell = B^{\ell-1} * B,$$

where the $*$ operator performs the correct computation of the vector values of each element in B^ℓ based on $B^{\ell-1}$ and B. Precisely, let $b^\ell_{i,j}$ denote the row-i column-j element in B^ℓ, then the $*$ operator performs the following:

$$b^\ell_{i,j} = \frac{1}{mn} \sum_{x=0}^{mn-1} b^{\ell-1}_{i,x} * b_{x,j},$$

where

$$(b^{\ell-1}_{i,x} * b_{x,j})[0] = 1 - (1 - b^{\ell-1}_{i,x}[0])(1 - b_{x,j}[0]),$$
$$(b^{\ell-1}_{i,x} * b_{x,j})[1] = 1 - (b^{\ell-1}_{i,x} * b_{x,j})[0] - (1 - b^{\ell-1}_{i,x}[0] - b^{\ell-1}_{i,x}[1])(1 - b_{x,j}[0] - b_{x,j}[1]),$$
$$\vdots$$
$$(b^{\ell-1}_{i,x} * b_{x,j})[t] = 1 - \sum_{k=0}^{t-1} (b^{\ell-1}_{i,x} * b_{x,j})[k] - (1 - \sum_{k=0}^{t} b^{\ell-1}_{i,x}[k])(1 - \sum_{k=0}^{t} b_{x,j}[k]),$$
$$\vdots$$
$$(b^{\ell-1}_{i,x} * b_{x,j})[N-1] = 1 - \sum_{k=0}^{N-2} (b^{\ell-1}_{i,x} * b_{x,j})[k].$$

To see why the above computation of $*$ is correct, we notice that the term $(b^{\ell-1}_{i,x} * b_{x,j})[t]$ stores the probability of the event that D_{S_t} is the shortest distance after ℓ trips when x is the ending point of the $(\ell-1)$-th trip. This event occurs if and only if the shortest distance attained at the first $\ell - 1$ trips and the shortest distance attained at the last trip are both at least D_{S_t}, but excluding the cases where the eventual shortest distance after ℓ trips is $D_{S_0}, D_{S_1}, \ldots$, or $D_{S_{t-1}}$. Based on this reasoning, we derive the formulation for $(b^{\ell-1}_{i,x} * b_{x,j})[t]$ as shown in the above definition.

C. Computing the theoretical lower bound

Once the matrix B^ℓ is computed, we can make use of the following theorem to obtain the expected shortest distance (which forms a lower bound on the expected communication distance).

Theorem 1 *The expected shortest distance to the target after ℓ trips can be calculated by:*

$$E[d_{min_\ell}] = \frac{1}{mn} \sum_{i=0}^{mn-1} \left(\frac{1}{mn} \sum_{j=0}^{mn-1} \sum_{k=0}^{N-1} D_{S_k} \times b^\ell_{i,j}[k] \right).$$

Proof: By the definition of B^ℓ, the expected shortest distance after ℓ trips, given cell i is starting location, is $(1/(mn)) \sum_{j=0}^{mn-1} \sum_{k=0}^{N-1} D_{S_k} \times b^\ell_{i,j}[k]$. The theorem thus follows since each cell i is equally likely to be the starting location in our map-based movement model. □

Theorem 1 illustrates the theoretical lower bound of the expected communication distance after a mobile node travels ℓ trips. However, in practice, it is more interesting to know the lower bound after a mobile node travels for some amount of *time* instead. E.g., we may want to know the lower bound expected communication distance if the mobile node is allowed to communicate at any time within the next 500 seconds. Thus, we may want to obtain the theoretical lower bound as a function of the total travel time, or more commonly, the *maximum allowable delay*.

In the following, we discuss how to obtain (or approximate) such a time-based lower bound under the rectified random waypoint model. Firstly, the lemma below gives the expected time for the mobile node to travel a single trip in the model.

Lemma 1 *Let $\alpha = \arctan(Y/X)$. In the rectified random waypoint model, the expected time for a single trip can be expressed by*

$$E[T] = \frac{\log(V_{max}/V_{min})}{V_{max} - V_{min}} \Bigg[\frac{X^3}{15Y^2}(1 - \frac{1}{\cos^3\alpha}) + \frac{X^2}{6Y}(\ln\frac{1+\sin\alpha}{\cos\alpha} + \frac{\sin\alpha}{\cos^2\alpha}) + \frac{Y^3}{15X^2}(1 - \frac{1}{\sin^3\alpha}) + \frac{Y^2}{6X}(\frac{\cos\alpha}{\sin^2\alpha} - \ln\frac{1-\cos\alpha}{\sin\alpha}) \Bigg].$$

Proof: Please refer to Appendix I. □

Based on the expected single trip time $E[T]$ in Lemma 1, when we want to obtain the lower bound of the expected communication distance after traveling some time t, we can use $\ell' = t/E[T]$ to estimate the number of trips traveled. In case ℓ' is an integer, we may apply Theorem 1 with $\ell = \ell'$ to obtain an approximation of the desired lower bound. However, in the general case where ℓ' is not an integer, we may then obtain this approximation through interpolation (of various $E[d_{min_\ell}]$ values computed by Theorem 1).

IV. Movement Prediction Algorithm

In previous sections, we illustrate how to obtain the theoretical lower bound of the expected communication distance. Such a theoretical lower bound provides the best achievable (i.e., shortest) communication distance between a mobile node and a receiver. Therefore, it remains for us to design an algorithm to achieve the communication distance as close to the theoretical lower bound as possible.

The design of such an algorithm has been studied in [1], in which several movement prediction algorithms are proposed and evaluated. Among them, the *least-distance* (LD) algorithm is the simplest and has the best practical performance.

The LD algorithm is based on the 37% rule of the Best-choice(r) algorithm, which solves the well known secretary problem [3]. In our case, we approximately track the first 37% or more of the candidate positions in the movement history of the mobile node, and find the least distance d_{min} between the mobile node and the receiver in the movement history. Then, in each of the next D time units, where D specifies the maximum allowable delay, we check if the current distance between the mobile node and the receiver is less than or equal to d_{min}. If so, we communicate immediately; otherwise, we communicate at the Dth time unit.

The performance of the LD algorithm *relative* to other prediction algorithms has been studied before [1]. In Section V, we will evaluate the effectiveness of the LD algorithm by comparing it with the theoretical lower bound of the expected communication distance, thus providing a new *absolute* measure of performance for the algorithm.

V. Simulation Results

This section verifies our theoretical results experimentally, and studies the performance of the LD algorithm under different network scenarios. The experiments are divided into three parts. Part *A* compares the lower bounds on the expected communication distance computed by the probability matrix approach in Section III and the approximation approach in Section VI. Part *B* illustrates the absolute performance of the LD algorithm. Both Part *A* and Part *B* assume the movement model to be the rectified random waypoint model. In Part *C*, we evaluate the performance of LD in the general map-based movement model. Further experimental results (e.g., the effects of the network cell size) can be found in [2].

A. Absolute performance of the LD algorithm

This section illustrates the absolute performance of the LD algorithm, by comparing the communication distance it achieves against the lower bound computed by the probability matrix approach and the lower bound obtained experimentally.

We again assume the rectified random waypoint model as the movement model. In the experiment, the mobile node moves in a 150 m by 150 m network area divided into 25 cells of size 30 m by 30 m. The target cell (inside which the receiver is located) is set to be the center of the network.

In Figure 1, we illustrate the results for different expected nodal speeds of 5 m/s, 10 m/s, 25 m/s, respectively by using the probability matrix approach. (The corresponding (V_{min}, V_{max}) for them are $(4, 6)$, $(10, 15)$, and $(20, 30)$, respectively.) We also the measured expected shortest distances and the average communication distance achieved by the LD algorithm, which are obtained experimentally over 100 independent 2,000,000-second simulation runs. We omit the error bars because the corresponding standard deviations are small.

From the figure, we observe that the performance of LD increases as the expected nodal speed increases or as the maximum allowable delay increases. We also notice that the calculated theoretical lower bound matches closely with the experimental lower bound. For the average reduction in the expected communication distance, we find that it is within 76% to 94% of the optimal.

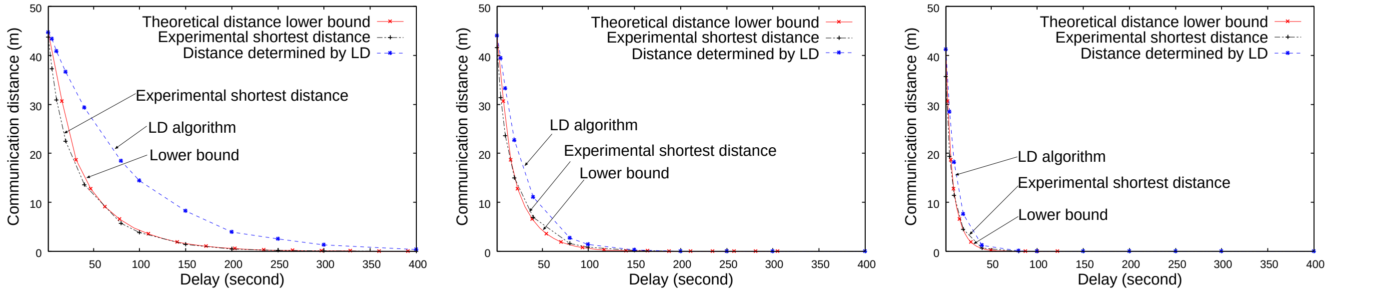
In Figure 2, we illustrate comparison results in a larger network. The dimensions of the network area are 270 m by 270 m, and the area is divided into 81 cells of size 30 m by 30 m; the target cell is set to be the center of the network. From the results, we find that with a larger network, it takes longer for a mobile node to find a closer position to communicate with the target. At the same time, the results show that our theoretical lower bound of the expected communication distance closely matches the experimental results. We find that the average reduction in the expected communication distance is within 75% to 94% of the optimal.

B. Performance of LD in map based movement model

In this subsection, we illustrate the performance of the LD algorithm in the general map-based movement model. In our model, we use a local map of Lafayette, Indiana, USA, as the accessibility constraints. The map is shown in Figure 3, which contains all the streets and roads in an 10,800 m by 11,500 m area. Ten mobile nodes are present in the network. The destinations of the nodes are chosen randomly and independently, and the nodes move along the routes generated by the route selection software. Each simulation runs 20,000 seconds. During each simulation run, each node communicates with a randomly chosen node as the receiver for 1000 times. We measure the distance of each communication, and report the average distance over three independent runs in Figure 4. Confidence intervals are also included in the graphs for reference.

Figure 4(a) shows the communication distance achieved by the LD algorithm as a function of the maximum allowable delay. We find that the LD algorithm significantly reduces the communication distance under the realistic instantiation of the map-based model in this section. The results show that LD has consistently good performance as in the prior work considering the simpler random waypoint movement model [1]. We find that the average reduction in the expected communication distance is around 62% of the optimal.

In Figure 4(b), we illustrate the percentage communication distance savings as a function of the maximum allowable delay. As the maximum allowable delay increases, the distance saving also increases. In the same graph, we show also the measured distance savings in relation to the theoretical upper

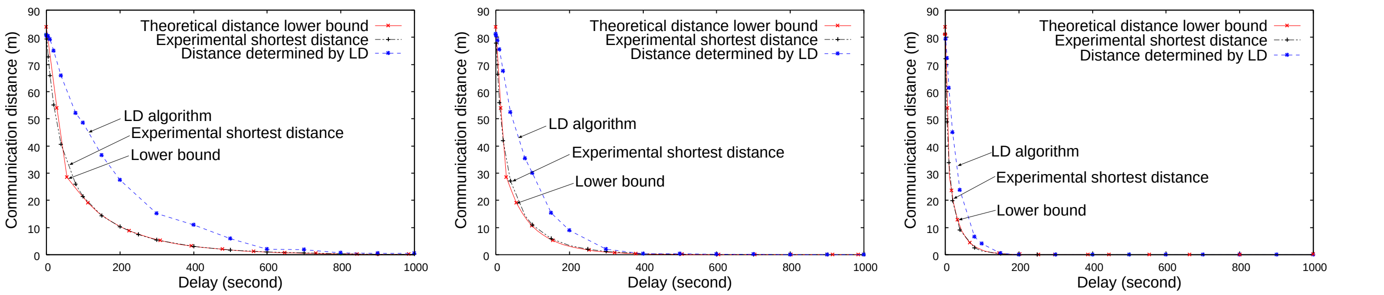


(a) Expected speed = 5 m/s

(b) Expected speed = 10 m/s

(c) Expected speed = 25 m/s

Fig. 1. Comparison of performance of LD with experimental expected shortest distance, and the theoretical lower bound in network area of size 150 m by 150 m divided into 25 cells. (Error bars are omitted because of small standard deviations.)

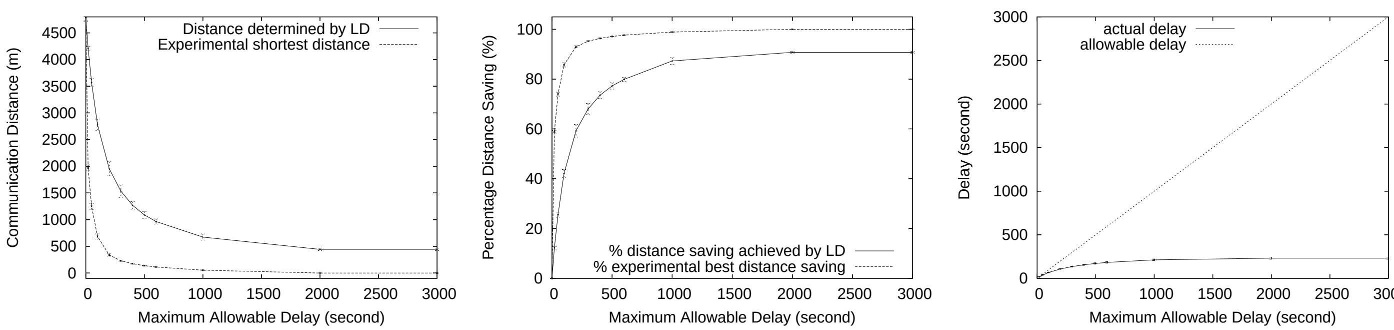


(a) Expected speed = 5 m/s

(b) Expected speed = 10 m/s

(c) Expected speed = 25 m/s

Fig. 2. Comparison of performance of LD with experimental expected shortest distance, and the theoretical lower bound in network area of size 270 m by 270 m divided into 81 cells. (Error bars are omitted because of small standard deviations.)



(a) Communication distance vs allowable delay

(b) Distance savings vs allowable delay

(c) Actual delay vs allowable delay

Fig. 4. Performance of LD with a map-based movement model: Communication distance is significantly reduced within relatively small delay.

bounds. Notice that in contrast to the random waypoint model results, the communication distances achieved by LD in this case do not converge to the experimental lower bounds, as the allowable delay increases. This is because the average speed is relatively low compared with the size of the area. (The average nodal speed is about 10 m/s due to speed limits of the local streets, whereas the area has dimensions 10,800 m by 11,500 m.) If a higher speed is allowed, we conjecture that the performance of LD would increase, according to observed trends under the rectified random waypoint model.

We also illustrate the actual communication delay by the LD algorithm in Figure 4(c), which is the time when LD decides to perform the communication. In general, LD will communicate before the allowable delay expires, when it concludes that the current position is likely to be close enough to the target receiver. By comparing the actual delay of LD with the allowable delay, we find that the algorithm is consistently able to predict a good position to communicate, within a small

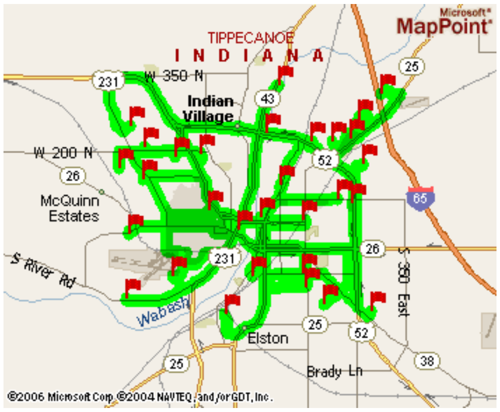


Fig. 3. A local map of Lafayette area.

fraction of the allowable delay.

To summarize, we have evaluated the LD algorithm in a realistic instantiation of the map-based movement model; we find that the LD algorithm continues to perform effectively and significantly reduce the communication distance.

C. Summary

By comparing the calculated theoretical lower bounds and the measured expected communication distances to the target cell, we can find that the theoretical bounds very closely match the measured results in all the network scenarios. Also, the LD algorithm shows good performance by achieving communication distances close to the lower bounds. Moreover, the performance improves as the nodal speed increases, verifying our claims in Section III.

The LD algorithm has been experimentally evaluated to perform well despite its simplicity, when compared with several other postponement algorithms in [1]. The results in this paper additionally quantify how close LD can match the theoretical lower bounds for any postponement algorithm. Moreover, LD's performance approaches the lower bounds as nodal speed increases.

VI. Further Discussion: An Approximation to the Theoretical Lower Bound

To calculate the theoretical lower bound for ℓ trips based on the probability matrix approach, we need to obtain B^ℓ first, which requires $O((mn)^3\ell)$ of the previously defined $*$ operations, each of which is performed on two length-N vectors. Thus, each $*$ operation takes $O(N)$ time, and the overall complexity is $O((mn)^3\ell N)$ time. This approach is slow even for moderate-size m and n. In this section, we give an alternative approach to obtain a close approximation of the theoretical lower bound of expected communication distance; the time complexity of the approximation approach is only $O(mn + N)$.

Our approximation approach considers the theoretical lower bound of expected communication distance as a function of the number of cells visited by the mobile node. It trades accuracy for speed by simplifying the calculations using two assumptions. In the following, we describe the assumptions the approximate approach makes, and then describe how to approximate the lower bound based on the number of cells visited.

A. Assumptions to simplify calculation

Let P_i denote the probability of a mobile node entering cell i when it leaves the current cell. Also, as in the previous section, we assume cells are partitioned into N sets according to their distances to the target cell, and we will re-use the notations of S_j's and D_{S_j}'s as before. For each set S_j, we define the probability P_{S_j} of a mobile node entering any cell in S_j when it leaves the current cell. Here, we have

$$P_{S_j} \approx \sum_{i \in S_j} P_i$$

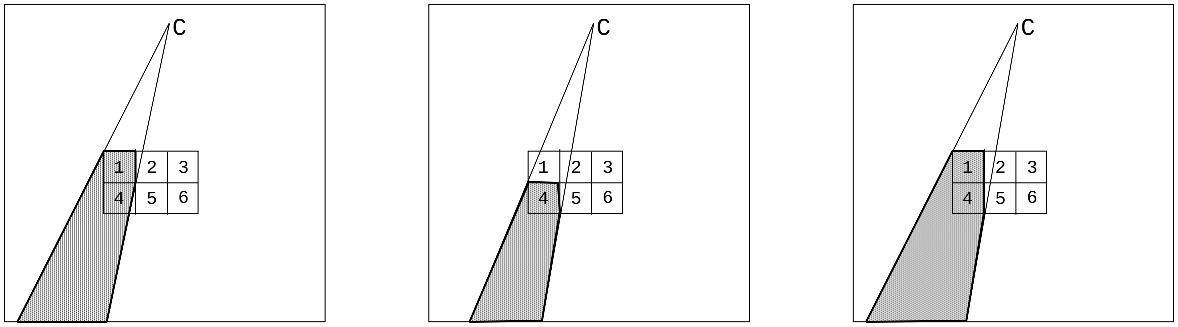
Note that the above equation is not an equality. For the first reason, it is because of the correlation among cells when a mobile node moves from one cell to another. To illustrate the correlation issue, we show that the probability that a mobile node visits a set of cells (i.e., to visit any cell in the set) is not the sum of the probabilities that it visits each individual cell in the set. In Figure 5, the point C indicates the current location of the mobile node. We want to find the probabilities for the mobile node to visit cell 1, to visit cell 4, or to visit either one of the cell, respectively.

Let the areas of the shaded regions in Figure 5(a)-5(c) be s_1, s_2, and s_3, respectively. Also, let the overall network area be S. Assuming the random waypoint movement model, the probability P_1 that the mobile node will visit cell 1 from location C is s_1/S. Similarly, the probability that the mobile node will visit cell 4 from C is given by $P_2 = s_2/S$. Finally, the probability of visiting either cell 1 or cell 4 is given by $P_3 = s_3/S$. We may find that $P_3 \neq P_1 + P_2$ as $s_1 + s_2 \neq s_3$, which means the probability that a mobile node moves to set of cells is not exactly equal to the sum of probabilities to move to each individual cell. Thus, in general, we cannot assume $P_{S_j} = \sum_{i \in S_j} P_i$. Nevertheless, as a first assumption of the approximation approach, we assume that the effect of such a correlation is negligible.

Even with the correlation issue neglected, we are still not able to assert $P_{S_j} = \sum_{i \in S_j} P_i$. The reason is that P_{S_j} should be dependent of the current location of the mobile node. Thus, our second assumption ssumes that irrespective of the current location, the mobile node will always move to a cell in S_j with probability $P_{S_j} = \sum_{i \in S_j} P_i$. Based on this, the next subsection gives the description of how we can approximate the theoretical lower bound efficiently.

We summarize our assumptions (though affecting the correctness, but simplifying the computation a lot) to get the approximated lower bound as follows:

- We assume that the correlation between a cat visiting each cell at the next step is negligible.
- We assume that the cat will always move to cell j next with a fixed probability P_{S_j}, irrespective of the current location.



(a) Probability for cell 1 (b) Probability for cell 4 (c) Probability for cells 1 and 4

Fig. 5. Example to show that probability of a set of cells is not exactly the sum of the probabilities of each cell in the set due to correlations among cells.

Later, in our simulation result, we shall see that the approximated lower bound is indeed very close to both the theoretical lower bound derived in Section III.

B. Approximating the theoretical lower bound

In this subsection, we based on the two assumptions and describe how to approximate the theoretical lower bound. Note that equality signs in the equations are in general incorrect, but will be correct under our two assumptions.

We define $P_{D_j}(k)$, where $0 \leq j \leq N-1$ to be the probability that the shortest distance between the mobile node and the target cell is D_{S_j} after visiting k cells. The calculation of $P_{D_j}(k)$ is then straightforward:

$$\begin{aligned}
P_{D_0}(k) &= P(\text{visiting some cell in } S_0) \\
&= 1 - P(\text{does not visit any cell in } S_0) \\
&= 1 - (1 - P_{S_0})^k, \\
P_{D_1}(k) &= P((\text{does not visit any cell in } S_0) \text{ and} \\
&\quad\; (\text{visiting some cell in } S_1)) \\
&= P(\text{does not visit any cell in } S_0) \cdot P((\text{visiting some} \\
&\quad\; \text{cell in } S_1) \mid (\text{does not visit any cell in } S_0))^1 \\
&= (1 - P_{S_0})^k \left[1 - (1 - \frac{P_{S_1}}{1 - P_{S_0}})^k \right], \\
P_{D_j}(k) &= (1 - P_{S_0} - P_{S_1} - ... - P_{S_{j-1}})^k \times \\
&\quad \left[1 - (1 - \frac{P_{S_j}}{1 - P_{S_0} - P_{S_1} - ... - P_{S_{j-1}}})^k \right], \\
&\vdots \\
P_{D_{N-1}}(k) &= P^k_{S_{N-1}}.
\end{aligned}$$

Then, we have the expected shortest distance as follows

$$E[d_g(k)] = \sum_{j=0}^{N-1} P_{D_j}(k) D_{S_j}.$$

The above equation gives the approximated lower bound of the expected communication distance as a function of number of cells a mobile node visited. Similar to the use of Lemma 1, we give the following lemma to help in expressing the approximated lower bound as a function of *time* instead.

[1] According to conditional probability, $P(AB) = P(A)P(B|A)$.

Lemma 2 *Within a single trip, the expected time the mobile node stay*

1) *in the starting and ending cells is*

$$\frac{8R}{3\pi(V_{max} - Vmin)} \log \frac{V_{max}}{V_{min}};$$

2) *in the cell when it is passing through the cell*

$$\frac{4R}{\pi(V_{max} - Vmin)} \log \frac{V_{max}}{V_{min}}.$$

Proof: Please refer to Appendix II. □

Suppose that we have obtained the probabilities P_{inside} and $P_{crossing}$ of a mobile node staying inside an end point and crossing a cell, respectively, through experiments. Then, we can calculate the expected time that a mobile node stays in a cell, called expected sojourn time $E[T_s]$, as

$$E[T_s] = P_{inside} E[T_i] + P_{crossing} E[T_c].$$

As an example of a network divided into 5 by 5 cells, P_{inside} and $P_{crossing}$ are obtained experimentally to be 0.454 and 0.546, respectively. Further discussions of the sojourn time of a mobile node can be found at [2].

By applying the expected cell sojourn time given by Lemma 2, we can calculate the expected shortest distance as a function of time t. Replacing k by the corresponding travel time (i.e., k multiplies the expected cell sojourn time $E[T_s]$ given by Lemma 2), we have the approximated theoretical lower bound of the expected communication distance as a function of the maximum allowable delay. Finally, for the complexity of the approximation approach, we observe that it takes $O(mn)$ time to compute all P_{S_j}'s, and $O(N)$ time to compute all $P_{D_j}(k)$'s for any k; in total, it takes $O(mn+N)$ time which is independent of k.

C. Simulation results: matrix approach versus approximation approach

This subsection compares the theoretical lower bounds derived by the probability matrix approach and the approximation approach. We assume the rectified random waypoint movement model with the following setting: the dimensions of the network area are 150 m by 150 m; the area is divided into 25 cells of size 30 m by 30 m; the target cell (inside which the receiver is located) is set to be the center of the network. We conduct three sets of calculations, assuming that the mobile node moves with an average speed of 5 m/s, 10 m/s, and 25 m/s, respectively. (The corresponding (V_{min}, V_{max}) for them are $(4, 6)$, $(10, 15)$, and $(20, 30)$, respectively.)

In Figure 6, we displayed the lower bounds as a function of the maximum allowable delay. From the figure, we conclude that though the lower bound from the approximation approach is not tightly matching the one computed by probability matrix approach, it is indeed a very close approximation.

We also report the times to compute the theoretical lower bound based on the two approaches. The maximum allowable delay is set to $40 \times E[T]$ (recall that $E[T]$ is expected time for a trip), which corresponds to the time for 40 trips in the probability matrix approach. Two set of timings are taken, for which the network is partitioned to 5×5 cells, and 9×9 cells, respectively. The results are presented in Table II. They show that the approximation approach takes significantly less time to compute than the exact approach.

TABLE II

TIME TO COMPUTE LOWER BOUND BY THE TWO APPROACHES. (DELAY = TIME FOR 40 TRIPS.)

number of cells	5×5	9×9
time for probability matrix approach (s)	0.174	10.287
time for approximation approach (s)	0.001	0.003

VII. Related Work

The random waypoint model is widely used in mobile networking research. Our map-based movement model is a significant generalization of the random waypoint model to incorporate (realistic) accessibility constraints and route selection algorithms according to the application. The movement of mobile users in a cellular network has been studied in [5], [8] to solve the handoff and channel allocation problems. Previous work on mobility prediction [4], [7] has been concerned with improving network connectivity, end-to-end delay, and network capacity. Their focus is thus quite different from ours, which is to understand how mobility can be used to provide a tradeoff between communication delay and communication distance. The random waypoint model has been analyzed in [9]. Their analysis focuses only on the steady state average nodal speed, which leads to the conclusion that the original model fails to sustain a meaningful average speed necessary for simulation experiments. Our analysis is significantly more comprehensive, and can be applied to understand the fundamental performance limits of movement predicted wireless communication. The expected sojourn time result in Lemma 2 (in Section VI-B) can be regarded as an extension of a similar result by Hong and Rappaport [5].

VIII. Conclusion

We have developed a realistic system model which allows general accessibility constraints and route selections to be specified for a mobile wireless network, while admitting the well known random waypoint model as a special case. For movement predicted communication under the random waypoint model, we have derived fundamental tight lower bounds of the expected communication distance achievable by *any* postponement algorithm, as a function of the sender's nodal speed and the allowable postponement delay. Our analysis provides an absolute performance measure of how closely a postponement algorithm can match the theoretical optimum. In particular, we show that the least-distance algorithm in prior work achieves a reduction in the expected communication distance within 62% to 94% of the optimal, over a realistic range of nodal speeds. Moreover, the algorithm's absolute performance increases as the nodal speed or the allowable delay increases. Our experimental results have further characterized the performance of movement prediction in an instantiation of the map-based model to road travel in the Lafayatte, Indiana area.

REFERENCES

[1] S. Chakraborty, Y. Dong, D. K. Y. Yau, and J. C. S. Lui, "On the Effectiveness of Movement Prediction to Reduce Energy Consumption in Wireless Communication," to appear in *IEEE Trans. on Mobile Computing.*

[2] Y. Dong, D. K. Y. Yau, and W. K. Hon, "Nodal Movement in Mobile Wireless Networks: Stochastic Analysis and Applications," Technical Report, CS Dept, Purdue University, 2006.

[3] P. R. Freeman, "The Secretary Problem and Its Extensions: A Review," *International Statistical Review,* 51, 1983.

[4] M. Grossglauser and D. Tse, "Mobility Increases the Capacity of Adhoc Wireless Networks," *IEEE/ACM Transactions on Networking,* 10(4), 2002.

[5] D. Hong and S. S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Non-prioritized Handoff Procedures," *IEEE Trans. on Vehicular Tech.,* 35(3), 1986.

[6] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks," *Ad Hoc Networks,* Chap. 5, 2001.

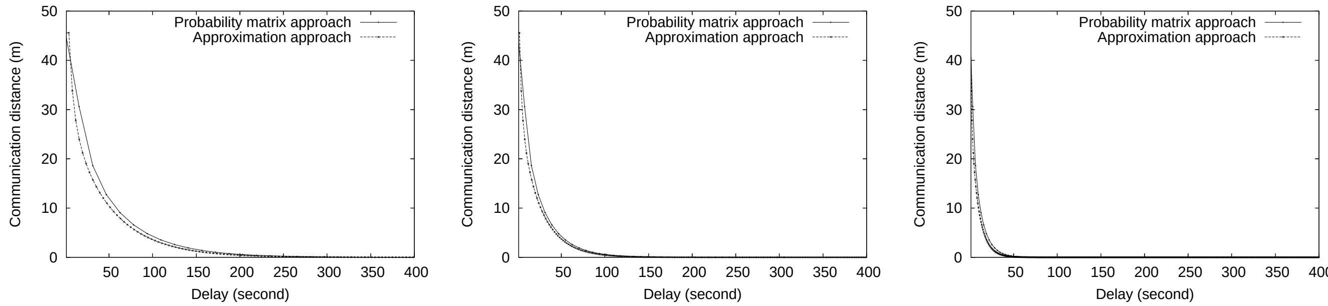
[7] W. Su, S. J. Lee, and M. Gerla, "Mobility Prediction and Routing in Ad Hoc Wireless Networks", *International Journal of Network Management,* 11(1), 2001.

[8] K. L. Yeung and S. Nanda, "Channel Management in Microcell/Macrocell Cellular Radio Systems," *IEEE Trans. on Vehicular Tech.,* 45(4), 1996.

[9] J. Yoon, M. Liu, and B. Noble, "Random Waypoint Considered Harmful," in *Proceedings of IEEE INFOCOM,* 2003.

APPENDIX I
Proof of Lemma 1

Proof: Let L denote the single trip distance. Let (x_1, y_1) and (x_2, y_2) denote the starting and ending points of a trip, respectively. Note that both the starting and destination points of a trip are uniformly chosen from within the network. Let $Z = |x_1 - x_2|$ and $W = |y_1 - y_2|$, so that the distance of the trip, L, is $\sqrt{Z^2 + W^2}$. To obtain the expected trip distance $E[L]$, we first compute the distributions of Z and W as follows.



(a) Expected speed = 5 m/s (b) Expected speed = 10 m/s (c) Expected speed = 25 m/s

Fig. 6. Lower bound of expected communication distance (as a function of delay) computed by probability matrix approach and approximation approach.

The pdf of x_1, or that of x_2, is given by $f(x) = 1/X$. The cdf of Z can then be calculated by:

$$\begin{aligned} F_Z(z) &= P(Z = |x_1 - x_2| < z) \\ &= \int_0^X \int_{x_2-z}^{x_2+z} f(x_1) f(x_2) dx_1 dx_2 \\ &= \int_{X-z}^{X+z} f(x_2) \int_{x_2-z}^{X} f(x_1) dx_1 dx_2 \\ &\quad + \int_z^{X-z} f(x_2) \int_{x_2-z}^{x_2+z} f(x_1) dx_1 dx_2 \\ &\quad + \int_0^z f(x_2) \int_0^{x_2+z} f(x_1) dx_1 dx_2 \\ &= \frac{2zX - z^2}{X^2}, \quad 0 \le z \le X. \end{aligned}$$

By differentiating $F_Z(z)$, we obtain the pdf of Z as follows:

$$f_Z(z) = F_Z'(z) = \frac{2}{X} - \frac{2z}{X^2}.$$

Similarly, the pdf of W is given by:

$$f_W(w) = F_W'(w) = \frac{2}{Y} - \frac{2w}{Y^2}.$$

Now, $E[L]$ can be calculated through the joint distribution of Z and W. Let $\alpha = \arctan(Y/X)$. We have

$$\begin{aligned} E[L] &= \int_0^Y \int_0^X \sqrt{z^2 + w^2} f_{Z,W}(z, w) dz dw \\ &= \int_0^Y \int_0^X \sqrt{z^2 + w^2} f_Z(z) f_W(w) dz dw \\ &= \int_0^Y \int_0^X \sqrt{z^2 + w^2} \left(\frac{2}{X} - \frac{2z}{X^2}\right)\left(\frac{2}{Y} - \frac{2w}{Y^2}\right) dz dw \\ &= \frac{X^2}{6Y}\left(\ln \frac{1 + \sin\alpha}{\cos\alpha} + \frac{\sin\alpha}{\cos^2\alpha}\right) + \frac{X^3}{15Y^2}\left(1 - \frac{1}{\cos^3\alpha}\right) \\ &\quad + \frac{Y^2}{6X}\left(\frac{\cos\alpha}{\sin^2\alpha} - \ln \frac{1 - \cos\alpha}{\sin\alpha}\right) + \frac{Y^3}{15X^2}\left(1 - \frac{1}{\sin^3\alpha}\right) \end{aligned} \tag{1}$$

For the expected time $E[T]$ of a single trip, since (i) $E[T] = E[LV^{-1}]$ and (ii) the random variables L and V are independent of each other, we have $E[T] = E[L]E[V^{-1}]$. For $E[V^{-1}]$, it can be calculated by:

$$\begin{aligned} E[V^{-1}] &= \int_{V_{min}}^{V_{max}} \frac{1}{v} f(v) dv = \int_{V_{min}}^{V_{max}} \frac{1}{v(V_{max} - V_{min})} dv \\ &= \frac{\ln(V_{max}/V_{min})}{V_{max} - V_{min}}. \end{aligned} \tag{2}$$

Lemma 1 thus follows by combining Equation 2 with Equation 1. □

APPENDIX II
Proof of Lemma 2

To begin with, we approximate the network cell by a circle that has the same area (see Figure 7). The radius of the circle, denoted by R, is therefore

$$R = \frac{s}{\sqrt{\pi}} \approx 0.56s.$$

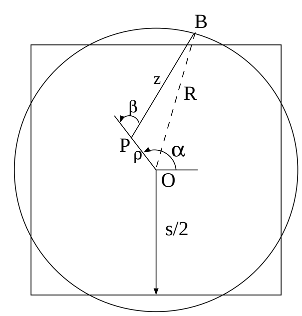


Fig. 7. Approximation of cell as a circle.

To calculate the expected cell sojourn time, we consider two scenarios where a sensor moves in a single trip: (1) Sensor starts or finishes a trip inside the cell; (2) Sensor passes through a cell during the trip. Lemma 2 gives the corresponding expressions.

Proof: We first calculate the expected sojourn time of Case (i). Let T_i denote the sojourn time that a mobile sensor is inside the starting cell, and let V denote the nodal speed. The joint pdf of T_i and V is given by

$$f_{T_i,V}(t,v) \quad = \quad |v| f_{Z,V}(z,v).$$

where Z is the distance from the sensor to the cell boundary.

The pdf of T_i can be calculated using the above joint pdf as follows: Let $C_1 = \frac{2}{\pi R^2(V_{max}-V_{min})}$ and $C_2(t) = \frac{8R}{3\pi t^2(V_{max}-V_{min})}$. Then, we have

$$f_{T_i}(t) = \int_{-\infty}^{\infty} f_{T_i,V}(t,v)dv = \int_{-\infty}^{\infty} |v| f_{Z,V}(z,v)dv$$

$$= \begin{cases} C_1 \int_{V_{min}}^{V_{max}} v\sqrt{R^2-(\frac{tv}{2})^2}dv, & 0 \le t \le \frac{2R}{V_{max}}; \\ C_1 \int_{V_{min}}^{2R/t} v\sqrt{R^2-(\frac{tv}{2})^2}dv, & \frac{2R}{V_{max}} \le t \le \frac{2R}{V_{min}}; \\ 0, & t \ge \frac{2R}{V_{min}}. \end{cases}$$

$$= \begin{cases} C_2(t) \left([1-(\frac{tV_{min}}{2R})^2]^{3/2} - [1-(\frac{tV_{max}}{2R})^2]^{3/2}\right), \\ \qquad\qquad 0 \le t \le \frac{2R}{V_{max}}; \\ C_2(t)\ [1-(\frac{tV_{min}}{2R})^2]^{3/2}, & \frac{2R}{V_{max}} \le t \le \frac{2R}{V_{min}}; \\ 0, & t \ge \frac{2R}{V_{min}}. \end{cases}$$

Then, the expected sojourn time in Case (i) becomes

$$E[T_i] = \int_{-\infty}^{\infty} t f_T(t)dt$$
$$= \frac{8R}{3\pi(V_{max}-V_{min})} \times$$
$$\left\{\int_0^{\frac{2R}{V_{max}}} \frac{1}{t}\left[(1-(\frac{tV_{min}}{2R})^2)^{3/2} - (1-(\frac{tV_{max}}{2R})^2)^{3/2}\right]dt \right.$$
$$\left. + \int_{\frac{2R}{V_{max}}}^{\frac{2R}{V_{min}}} \frac{1}{t}(1-(\frac{tV_{min}}{2R})^2)^{3/2}dt\right\}$$
$$= \frac{8R}{3\pi(V_{max}-V_{min})} \ln\frac{V_{max}}{V_{min}}.$$

To calculate the expected sojourn time in Case (ii), let T_c denote the sojourn time when the sensor is in a cell other than the starting and the ending cells (i.e., when the sensor is crossing a cell in the trip). Let Z_c be the length of chord where the path of the sensor intersects the cell. The pdf of Z_c is

$$f_{Z_c}(z) = \frac{2}{\pi\sqrt{(2R)^2+z^2}}, \quad 0 \le z \le 2R.$$

Then, the expected chord length is

$$E[Z_c] = \int_0^{2R} z f_Z(z)dz = \frac{4R}{\pi}.$$

Also, in the proof of Lemma 1, we know that

$$E[V^{-1}] = \frac{\ln(V_{max}/V_{min})}{V_{max}-V_{min}}.$$

Then, the expected sojourn time in Case (ii) becomes

$$E[T_c] = E[Z_c]E[V^{-1}] = \frac{4R}{\pi(V_{max}-V_{min})}\ln\frac{V_{max}}{V_{min}}.$$

This completes the proof of Lemma 2. □