On Hop Count and Euclidean Distance in Greedy Forwarding in Wireless Ad Hoc Networks

Swades De, *Member, IEEE*

***Abstract*— In this letter, a probabilistic analysis is presented that captures the bounds on hop count from a given Euclidean distance between two nodes and vice versa in a greedy forwarding in wireless ad hoc networks. Accuracy of the analysis is verified via network simulations. The results could be useful in ad hoc and sensor network design and performance evaluation.**

Index Terms – ad hoc networks, sensor networks, greedy routing, probabilistic bounds

I. Introduction

Wireless ad hoc networks are generally characterized by randomly located nodes and peer-to-peer multihop connections. Geographic location aware greedy routing is one of the widely studied ad hoc routing protocols [1],[2]. Although greedy routing may not guarantee minimum energy or high capacity routes, its importance lies in its simplicity and scalability. Moreover, in miniaturized network nodes, such as sensors, processing overhead is comparable to, or even higher than, the transmission related power consumption [3], and due to limited power saving and added protocol complexity, transmit power control may not be cost-effective. As a result, in ad hoc network applications greedy forwarding may still remain an important routing candidate. The random nature of node distribution implies that for a given source-to-destination Euclidean distance, the number of hops is non-deterministic, and similarly, the Euclidean distance coverage in a given number of hops is not fixed. A few recent papers (e.g., [4],[5]) addressed the problem of establishing the relationship between Euclidean distance and hop count. However, the evaluation of tight bounds relating these distances has not been reported yet.

In this paper, the relationship of hop count and Euclidean distance between two nodes is analyzed. A greedy routing approach - called *least remaining distance (LRD) forwarding* - is considered that attempts to minimize the remaining distance to the destination in each hop. LRD forwarding is different from the maximum forwarding with fixed radius (MFR) [6],[7],[8], and the maximum forwarding with variable radius (MVR) [8] approaches. In particular, as observed in [7], although MFR and MVR ensure the maximum progress towards the destination, they do not guarantee minimizing the remaining distance to the destination. The LRD forwarding captures the cases for nodes with fixed as well as variable transmission radius, i.e., without and with transmission power control.

Manuscript received May 24, 2005.

S. De is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA (email: swadesd@njit.edu).

II. Analysis on Bounds

In the following analysis, the nodes are considered uniformly randomly placed in a two-dimensional space. All nodes have equal, omnidirectional transmission pattern of range R. As in [8], LRD approach tries to avoid "backward movement" if no forwarding node closer to the destination is found. For simplicity of the analysis, a node is considered to be a potential forwarder if it is in the half circle of the transmission range of a node towards the destination (the entire shaded region in Fig. 1). Precisely, this approach does not guarantee that the remaining distance would be always lesser than the current distance. Referring to Fig. 1, if the selected forwarding node P

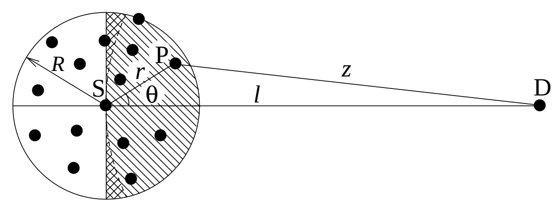

Fig. 1. In the LRD approach, a potential forwarding node P can be located anywhere inside the shaded region.

is located in the densely shaded region, the remaining distance to the destination D would be larger than the current distance between S and D. However, via simple geometry it can be shown that with a reasonably high node density the potential of backward movement is very insignificant – even when the densely shaded region is included. For example, with average 10 neighbors (average number of nodes within the coverage region of a node), the maximum probability of backward movement is 0.013, whereas with average 20 neighbors, this probability is 3.56×10^{-4}. Note that the average number of neighbors required to ensure only one-connectivity was shown in [7] to be 8.

A. Least remaining distance (LRD)

Denote the distance between a node S and the destination D by l. Let P be a potential forwarder of S, randomly located at a distance $\boldsymbol{r}$ and angle $\boldsymbol{\theta}$, and let the remaining distance from P to D be $\boldsymbol{z}$. With respect to the node S, the random position of P is characterized by the following joint probability density function (pdf) of the random variables (RVs) $\boldsymbol{r}$ and $\boldsymbol{\theta}$:

$$f_{\boldsymbol{r\theta}}(r,\theta) = \begin{cases} \frac{2r}{\pi R^2}, & 0 \le r \le R \text{ and } -\frac{\pi}{2} \le \theta \le \frac{\pi}{2} \\ 0, & \text{elsewhere.} \end{cases} \quad (1)$$

From geometry, $\boldsymbol{z} = \sqrt{(l-\boldsymbol{x})^2 + \boldsymbol{y}^2}$, where $\boldsymbol{x} = \boldsymbol{r}\cos\boldsymbol{\theta}$ and $\boldsymbol{y} = \boldsymbol{r}\sin\boldsymbol{\theta}$. Using the method of transformation of variables,

$$f_{\boldsymbol{z}}(z)=\begin{cases}\frac{4z}{\pi R^2}\arccos\left(\frac{l^2+z^2-R^2}{2lz}\right), & l-R\le z\le l\\ \frac{4z}{\pi R^2}\left[\arcsin\left(\frac{l}{z}\right)-\arcsin\left(\frac{l^2+z^2-R^2}{2lz}\right)\right], & l\le z\le\sqrt{l^2+R^2}\\ 0, & \text{elsewhere.}\end{cases} \tag{2}$$

we have the pdf of $\boldsymbol{z}$ given in (2). The corresponding cumulative distribution function (cdf) is $F_{\boldsymbol{z}}(z)=\int_{-\infty}^{z}f_{\boldsymbol{z}}(t)dt$.

Let the remaining distances from $\boldsymbol{n}$ potential forwarding nodes be $\boldsymbol{z}_1,\ \boldsymbol{z}_2,\ \cdots,\boldsymbol{z}_{\boldsymbol{n}}$, where $\boldsymbol{n}$ is a RV. By Poisson approximation of uniformly random node distribution,

$$\Pr[\boldsymbol{n}=n]\triangleq P(n)=\frac{(\rho a)^n}{n!}e^{-\rho a} \tag{3}$$

where ρ is the node density and $a=\pi R^2/2$. The average of $\boldsymbol{n}$ is $\overline{\boldsymbol{n}}=\rho a$. In each hop, the next forwarding node P to D has the least remaining distance $\boldsymbol{\xi}$, that satisfies the relation: $\boldsymbol{\xi}\ =\ \min\left\{\boldsymbol{z}_1,\ \boldsymbol{z}_2,\ \cdots,\ \boldsymbol{z}_{\boldsymbol{n}}|\boldsymbol{n}=n\right\}$. Since the nodes are uniformly random distributed, $\boldsymbol{z}_i\ \forall\ i=1$ to $\boldsymbol{n}$ are independent and identically distributed (iid) RVs. Moreover, since $\boldsymbol{n}$ is also a RV with the distribution given in (3), by Bayes' theorem,

$$f_{\boldsymbol{\xi}}(\cdot)=\sum_{n=0}^{\infty}P(n)nf_{\boldsymbol{z}}(\cdot)R_{\boldsymbol{z}}^{n-1}(\cdot) \tag{4}$$

where $f_{\boldsymbol{z}}(z)$ is given in (2), and $R_{\boldsymbol{z}}=1-F_{\boldsymbol{z}}$.

B. One hop progress

With the known characteristic of least remaining distance $\boldsymbol{\xi}$, the maximum forward progress $\boldsymbol{\varepsilon}$ can be obtained from the relation $\boldsymbol{\varepsilon}=l-\boldsymbol{\xi}$. The pdf of $\boldsymbol{\varepsilon}$ is $f_{\boldsymbol{\varepsilon}}(\varepsilon)=f_{\boldsymbol{\xi}}(l-\varepsilon)$, and the cdf is $F_{\boldsymbol{\varepsilon}}(\varepsilon)=1-F_{\boldsymbol{\xi}}(l-\varepsilon)$. Using (2)–(4) and simplifying, we get the exact expression of $f_{\boldsymbol{\varepsilon}}(\varepsilon)$. It can be easily verified that at moderately high node density (e.g., $\overline{\boldsymbol{n}}\ge 8$), $f_{\boldsymbol{\varepsilon}}(\varepsilon)\approx 0$, for $l-\sqrt{l^2+R^2}\le\varepsilon\le 0$ (also see Fig. 2). Accordingly, the

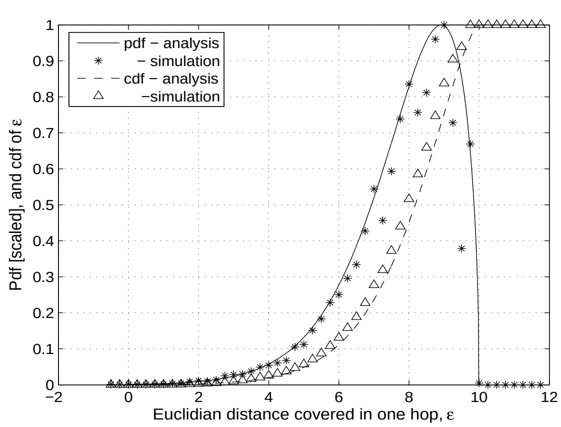

Fig. 2. Pdf of distance progress in one hop. $\overline{\boldsymbol{n}}=8$, $R=10$, $l=100$.

approximate expression of the pdf $f_{\boldsymbol{\varepsilon}}(\cdot)$ is shown in (5).

The k-th moment of $\boldsymbol{\varepsilon}$ can be calculated from $f_{\boldsymbol{\varepsilon}}(\cdot)$ via numerical integration, $\overline{\boldsymbol{\varepsilon}^k}\ =\ \int_0^R\varepsilon^k f_{\boldsymbol{\varepsilon}}(\varepsilon)d\varepsilon$, from where the mean $\overline{\boldsymbol{\varepsilon}}$ and standard deviation $\sigma=\sqrt{\overline{\boldsymbol{\varepsilon}^2}-(\overline{\boldsymbol{\varepsilon}})^2}$ are obtained.

TABLE I
MEAN ($\overline{\boldsymbol{\varepsilon}}$) AND STANDARD DEVIATION (σ) OF ONE-HOP PROGRESS AS A FUNCTION OF NODE DENSITY AND DISTANCE.

	$\overline{\boldsymbol{n}}=5$		$\overline{\boldsymbol{n}}=20$		$\overline{\boldsymbol{n}}=35$	
l	$\overline{\boldsymbol{\varepsilon}}$	σ	$\overline{\boldsymbol{\varepsilon}}$	σ	$\overline{\boldsymbol{\varepsilon}}$	σ
50	6.982	2.074	8.82	0.812	9.19	0.565
100	7.066	2.037	8.86	0.788	9.217	0.547
200	7.106	2.018	8.878	0.776	9.231	0.539

As shown in Table I, the mean ($\overline{\boldsymbol{\varepsilon}}$) and standard deviation (σ) of one-hop progress are slowly varying functions of distance to the destination, and can be considered approximately constants, estimated based on the expected distance between a source-destination pair in a given network area. Hence, for a given Euclidean distance l, the approximate average hop count is $\overline{h}=\lceil\frac{l}{\overline{\boldsymbol{\varepsilon}}}\rceil$. The bounds on hop count for a given Euclidean distance l can be numerically computed from (5).

C. Bounds on distance coverage

Let us denote the total number of hops between two end nodes as $h+1$. Observe that the LRD forwarding approach is active in first h hops, and the distance coverage in the last hop to the destination is uniformly distributed in $(0,R]$. While the last hop statistics can be easily accommodated in the bounds, for the sake of concise presentation of the LRD forwarding, henceforth only first h-hop statistics will be considered.

Assume that the forward progress in each hop is independent of the previous hops. Then, the total progress in h hops is a RV: $\boldsymbol{\varepsilon}^{'}=\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2+\cdots+\boldsymbol{\varepsilon}_h$, where $\boldsymbol{\varepsilon}_i,\ \forall\ i=1$ to h, are iid RVs with mean $\overline{\boldsymbol{\varepsilon}}$ and variance σ^2. For moderately large network (with $h>3$), by central limit theorem, $\boldsymbol{\varepsilon}^{'}$ is normal distributed with mean $h\overline{\boldsymbol{\varepsilon}}$ and variance $h\sigma^2$. Since $\boldsymbol{\varepsilon}^{'}$ is normal distributed, a proper multiplication factor k can be chosen such that the distance coverage in first h hops always lies within the region $k\sqrt{h}\sigma$ around the mean value with very high probability. In other words, by proper choice of k it can be ensured that the distance covered in h hops lies between

$$d_l=h\overline{\boldsymbol{\varepsilon}}-k\sqrt{h}\sigma\ \ \text{and}\ \ d_u=h\overline{\boldsymbol{\varepsilon}}+k\sqrt{h}\sigma. \tag{6}$$

For e.g., with $k=4$, probability that the distance covered in h hops is beyond $h\overline{\boldsymbol{\varepsilon}}\pm k\sqrt{h}\sigma$ is nearly 10^{-5}.

The limiting case is studied to verify the intuition that the difference between bounds on distance coverage d_u-d_l tends to zero as the node density increases. From (5), we have the cdf of $\boldsymbol{\varepsilon}$ in (7). First, note from (3) that at large node density (i.e., when $\overline{\boldsymbol{n}}\ =\ \rho a\ \gg\ 1$), the probability $P(n)\to 0$ and hence $F_{\boldsymbol{\varepsilon}}(\varepsilon)\to 0$, for small n. To compute $F_{\boldsymbol{\varepsilon}}(\varepsilon)$ for large n (i.e., when $P(n)$ is non-zero), denote $x=\frac{2}{\pi R^2}\left\{\frac{1}{2}\sqrt{4R^2l^2-(R^2-\varepsilon^2+2l\varepsilon)^2}-(l-\varepsilon)^2\arccos\left(1+\frac{\varepsilon^2-R^2}{2l(l-\varepsilon)}\right)+R^2\arcsin\left(\frac{R^2-\varepsilon^2+2l\varepsilon}{2lR}\right)\right\}$. We have, $0\le x\le 1$

$$f_{\boldsymbol{\varepsilon}}(\varepsilon)=\begin{cases}\sum_{n=0}^{\infty}P(n)2n\left(\frac{2}{\pi R^2}\right)^n(l-\varepsilon)\arccos\left(1+\frac{\varepsilon^2-R^2}{2l(l-\varepsilon)}\right)\left[\frac{1}{2}\sqrt{4R^2l^2-(R^2-\varepsilon^2+2l\varepsilon)^2}\right. & \\ \quad\left.-(l-\varepsilon)^2\arccos\left(1+\frac{\varepsilon^2-R^2}{2l(l-\varepsilon)}\right)+R^2\arcsin\left(\frac{R^2-\varepsilon^2+2l\varepsilon}{2lR}\right)\right]^{n-1}, & 0\le\varepsilon\le R\\ 0, & \text{elsewhere.}\end{cases}\tag{5}$$

$$F_{\boldsymbol{\varepsilon}}(\varepsilon)=\begin{cases}0, & \varepsilon<0\\ \sum_{n=0}^{\infty}P(n)\left[\frac{2}{\pi R^2}\left\{\frac{1}{2}\sqrt{4R^2l^2-(R^2-\varepsilon^2+2l\varepsilon)^2}-(l-\varepsilon)^2\arccos\left(1+\frac{\varepsilon^2-R^2}{2l(l-\varepsilon)}\right)+R^2\arcsin\left(\frac{R^2-\varepsilon^2+2l\varepsilon}{2lR}\right)\right\}\right]^n, & 0\le\varepsilon\le R\\ 1, & \varepsilon>R.\end{cases}\tag{7}$$

when $0\le\varepsilon\le R$. By the property of cdf, $\forall\ 0\le\varepsilon<R$ (i.e., $0\le x<1$), $F_{\boldsymbol{\varepsilon}}(\varepsilon)<1$, and $\forall\ \varepsilon\ge R$, $F_{\boldsymbol{\varepsilon}}(\varepsilon)=1$. Since $\lim_{n\to\infty}x^n=0\ \forall\ x<1$, the cdf $F_{\boldsymbol{\varepsilon}}(\varepsilon)$ is a step function at large n, i.e., at $n\to\infty$, $F_{\boldsymbol{\varepsilon}}(\varepsilon)=0\ \forall\ \varepsilon<R$ and $F_{\boldsymbol{\varepsilon}}(\varepsilon)=1\ \forall\ \varepsilon\ge R$. Thus, in the limit, the pdf $f_{\boldsymbol{\varepsilon}}(\varepsilon)=\delta(\varepsilon-R)$, a shifted delta function, i.e.,

$$\lim_{n\to\infty}f_{\boldsymbol{\varepsilon}}(\varepsilon)=\begin{cases}1 & \text{at }\varepsilon=R\\ 0 & \text{elsewhere,}\end{cases}$$

which implies that at infinitely large node density, $\boldsymbol{\varepsilon}\to R$ (constant), and the standard deviation of $\boldsymbol{\varepsilon}$, $\sigma\to0$. Hence from (6), for a finite network size (i.e., finite h) the difference, $d_u-d_l\to0$.

III. Results

The analytic and numerically computed bounds are verified via network simulations in C. In the simulations, nodes are uniformly random distributed in a 400×400 location space. The transmission range of a node is fixed at $R=10$. The total number of nodes is varied appropriately to attain a desired node density $\rho=\frac{2\overline{\boldsymbol{n}}}{\pi R^2}$. To minimize the boundary effect, the end nodes are chosen along the diagonals of the rectangular space, at least unit range inside the edges. In Fig. 2, the analytic distributions of one-hop progress are verified by simulation. Fig. 3 shows the average and bounds on Euclidean distance at different hop counts. Bounds from network simulations show very good match with the analysis (cf. (6)), and the results are verified to hold at any node density that ensures network connectivity. Also note the tighter upper bound of distance coverage compared to the deterministic upper bound hR. Numerically computed hop bound results are also verified to match well with the simulation.

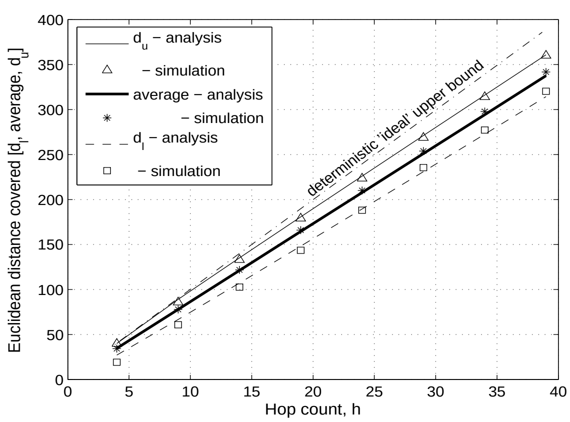

Fig. 3. Bounds on Euclidean distance. $\overline{\boldsymbol{n}}=15$, $R=10$.

IV. Concluding Remarks

In this letter, the relationship of Euclidean distance and hop count in greedy *least remaining distance* forwarding was characterized. It was shown that the moments of one-hop distance progress are slowly varying functions of the distance to the destination, and can be considered nearly constant for a given node density and network size. From these results, the average hop count was obtained for a given Euclidean distance. Further, from hop count, the average and bounds of Euclidean distance were derived. Network simulations verified correctness of the analysis.

The bounds on Euclidean distance and hop count could be useful in many ad hoc network applications, such as estimating end-to-end delay and jitter, total power consumption along the route, relative distance from hop counts in GPS-less positioning and routing approaches [4],[9]. In practice, the effects of network congestion, residual power awareness, wireless channel conditions, etc. can be added on top of these basic estimates.

References

[1] B. Karp and H. T. Kung, “GPSR: Greedy perimeter stateless routing for wireless networks,” in *Proc. ACM/IEEE MobiCom*, Boston, MA, Aug. 2000, pp. 243–254.

[2] M. Mauve, J. Widmer, and H. Hartenstein. “A survey on position-based routing in mobile ad hoc networks,” *IEEE Network Magazine*, vol. 15, no. 6, pp. 30–39, June 2001.

[3] *http://www.chipcon.com/files/CC2420_Data_Sheet_1_2.pdf*, June 2004.

[4] R. Nagpal, H. Shrobe, and J. Bachrach, “Organizing a global coordinate system from local information on an ad hoc sensor network,” in *Proc. Workshop on Information Processing in Sensor Networks*, Palo Alto, Apr. 2003, Springer LNCS, vol. 2634, pp.333–348.

[5] R. Hekmat and P. V. Mieghem, “Degree distribution and hop count in wireless ad hoc networks,” in *Proc. IEEE ICON*, Sydney, Australia, Sept. 2003, pp. 603–609.

[6] L. Kleinrock and J. Silvester, “Optimum transmission radii for packet radio networks or why six is a magic number,” in *Proc. IEEE National Telecomm. Conf.*,, Dec. 1978, pp. 4.3.1–4.3.5.

[7] H. Takagi and L. Kleinrock, “Optimal transmission ranges for randomly distributed packet radio terminals,” *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 246–257, Mar. 1984.

[8] T.-C. Hou and V. O. K. Li, “Transmission range control in multihop packet radio networks,” *IEEE Trans. Commun.*, vol. 34, no. 1, pp. 38–44, Mar. 1986.

[9] A. Rao, C. Papadimitrou, S. Ratnasamy, S. Shenker, and I. Stoica, “Geographic routing without location information,” in *Proc. ACM/IEEE MobiCom*, San Diego, CA, Sept. 2003, pp. 96–108.