Meeting Connectivity Requirements in a Wireless Multihop Network

Ji Li, *Student Member, IEEE*, Lachlan L. H. Andrew, *Member, IEEE*, Chuan Heng Foh, *Member, IEEE*, Moshe Zukerman, *Senior Member, IEEE*, and Marcel F. Neuts

Abstract— This paper investigates the connectivity probability of 1-dimensional ad hoc networks in which nodes have random, non-identically distributed locations, this leads to optimization of the number of nodes required. An empirical approach is used. We fit a parametric distribution to the CDF of the maximum distance between adjacent nodes. Special and extreme cases which are not covered by the empirical approach are treated separately.

Index Terms— Ad hoc network, sensor network, connectivity probability, empirical analysis, log-logistic distribution.

I. INTRODUCTION

AD HOC and sensor networks are increasingly drawing attention due to the variety of potential applications and advances in wireless technologies. Sensor networks aim to achieve high reliability, flexible utilization, cost-effectiveness and ease of deployment [1]. Although it is very important to optimize the placement of sensors, it is often difficult to do that; e.g., it is difficult to find a suitable mounting point exactly at the target location, causing it to be displaced by a random distance, or our sensor may move after placement due to winds, floods, etc. Accordingly, randomness in ad hoc networks is unavoidable in many cases.

It is important to ensure that such random networks are connected; that is, there is a path from each node to each other node, using hops no larger than the radio transmission range. Gupta and Kumar [2] used connectivity requirements to bound the capacity of wireless networks. Bettstetter investigated the number of nodes needed to maintain k-connectivity in a certain area [3] and the effects of shadowing upon connectivity [4]. Foh *et al.* derived formulas to calculate the network connectivity for uniformly and non-uniform distributed sensors in [5] and [6], respectively.

These studies have assumed that the locations of nodes are identically distributed. We consider the case of a one dimensional point process, with N points randomly distributed

around uniformly spaced mean locations. For example, this may arise if nodes are dropped at equal intervals from a vehicle (either airborne or terrestrial) traveling at a constant speed. In particular, we aim to establish an ad hoc network between a source and destination at known locations.

The question whether the network is connected is mathematically equivalent to testing the hypothesis that all physically adjacent sensors are within a certain distance from each other. This gives an obvious sufficient statistic, the maximum distance between two physically adjacent nodes. Its probability distribution is crucial to the application of that hypothesis test or equivalently to the network connectivity probability. That distribution is easily described in terms of the underlying events in the experiment sample space, but the analytic expression involves a complicated N-fold integral which defies explicit evaluation.

The estimation of this probability distribution is readily performed by simulating a large number of N-tuples of locations of the sending nodes. We empirically investigate the probability of connectivity of this network in Section III by fitting a curve to the Cumulative Distribution Function (CDF) of the maximum distance between adjacent nodes. The parameters of this curve are expressed as empirical functions of the problem parameters. Curve fitting using experimental or simulation results is a very common technique in many fields of science and engineering. It is, however, not very common in networking research. One example of curve fitting application in network research is the OSPF performance analysis by Cui *et al.* [7]. In Section IV, we consider some special cases in which our fitting procedure is not applicable.

II. MODEL DESCRIPTION

The network seeks to connect a source node to a destination, using N intermediate nodes. Two nodes can communicate directly if and only if the distance between them is less than or equal to the radio range d. We consider a one dimensional coordinate system with the source at the origin and the destination at point $D = 1$. The setting of $D = 1$ is without loss of generality as it represents a scaling of all spatial variables. The location of the ith intermediate node is assumed to follow a Gaussian distribution with mean $\mu_i = i/(N+1)$ and variance σ^2 independent of all other nodes, as shown in Fig. 1. Let x_i, $i = 1, 2, ..., N$ denote the location of the ith sensor. Note that the locations of the sensors may not be in the same order as they are dropped. That is, we may have $x_i \geq x_{i+1}$ for some i. Let $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(N)}$ be
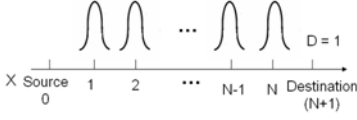
Fig. 1.   Sensor placement model

Fig. 2.   Curve fitting for CDF of d_m

a relabeling of the node locations in increasing order based on their physical locations. To investigate the probability of connectivity of a given network, we studied the CDF of the maximum distance between physically adjacent sensors $\max_i(x_{(i+1)} - x_{(i)})$ by simulation. For a given sensor network, let C be the event that the network is connected, denote the probability of connectivity as $P(C|d, N, \sigma)$, and let $d_m = \max_i(x_{(i+1)} - x_{(i)})$. Note that the probability of connectivity is a function of d, N, and σ. The network is connected if and only if $d_m \leq d$. Therefore:

$$P(C|d, N, \sigma) = P(d_m \leq d). \tag{1}$$

Let N^* be the minimal value of N, which satisfies the constraint: $P(C|N, \sigma, d) \geq P(target)$; i.e., N^* is the smallest number of sensors required for our one dimensional network to be connected with a pre-specified probability $P(target)$. We aim to find a good approximation for N^*.

III. AN APPROXIMATION FOR N^*

We empirically obtain the CDF of d_m parameterized by N and σ in two steps: 1) for many pairs (N, σ) the CDF of d_m is modeled by a parametric distribution; 2) the parameters of these distributions are expressed as functions of N and σ which leads to an approximate CDF of d_m.

A. Step 1:

We considered $38 \times 26 = 988$ pairs (N, σ) with N = 10, 20, ..., 190, 200, 300, ... 2000; σ = 0.005, 0.010,..., 0.045, 0.05, 0.10,..., 0.85. For each of these (N, σ) pairs we obtained 10^6 values of d_m by Monte Carlo simulation yielding an empirical CDF. As the example illustrated in Fig. 2, these CDFs are well approximated by a Log-Logistic distribution [8]:

$$P(d_m < x) = \frac{1}{1 + \exp[-(\log x - \alpha)/\beta]}. \tag{2}$$

This also holds for N as low as 10. For example, when $N = 10$ and $\sigma = 0.15$, the maximum deviation of the fitted Log-logistic CDF from the true CDF is only 2.1%.

The method we selected to fit the Log-Logistic function is Maximum Likelihood Estimation [8], [9]. For each pair (N, σ), we can get a pair of parameters α and β.

B. Step 2:

We fit α and β as functions of the parameters (N, σ), such that we can estimate N from (2) with given P, σ and d.

No simple curves were found which fit these surfaces well for all N and σ. In particular, there are difficulties in cases of small values for N and σ which will be treated separately in the next section. When $N \geq 10$ and $\sigma \geq 0.005$, the surface can be well approximated by piecewise functions, with boundaries
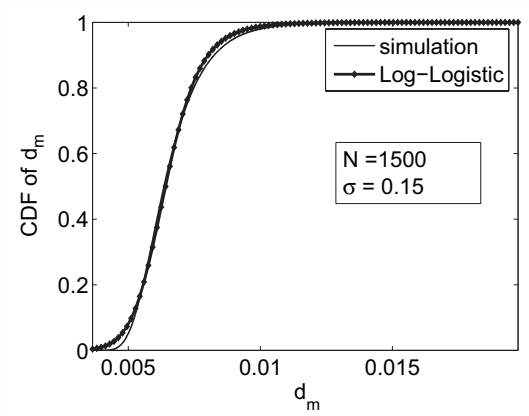
at $N = 200$ and $\sigma = 0.1$. The four pieces will be denoted xy, where $x = 0$ if $N < 200$ and $x = 1$ otherwise, and $y = 0$ if $\sigma < 0.1$ and $y = 1$ otherwise. Then, α and β are given by:

$$\alpha = \begin{cases} a_{11} + a_{12} \log N + a_{13}/\sqrt{\sigma} & \text{in piece 00} \\ a_{21} + a_{22} \log N + a_{23}\sigma & \text{in piece 01} \\ a_{31} + a_{32} \log N + a_{33}(\log \sigma)^2 & \text{in piece 10} \\ a_{41} + a_{42} \log N + a_{43}\sigma & \text{in piece 11} \end{cases} \tag{3}$$

$$\beta = \begin{cases} b_{11} + b_{12} \log N/N + b_{13}\sqrt{\sigma} \log \sigma & \text{in pieces 00} \\ 1/(b_{21} + b_{22}(\log N)^2 + b_{23} \log \sigma) & \text{in pieces 01} \\ 1/(b_{31} + b_{32}\sqrt{N} \log N + b_{33}/\sqrt{\sigma}) & \text{in pieces 10} \\ b_{41} + b_{42} \log N + b_{43}\sigma^2 \log \sigma & \text{in pieces 11} \end{cases} \tag{4}$$

where a_{ij} and b_{ij} are elements of matrix A and B, respectively:

$$A = \begin{bmatrix} -0.00732 & -0.12021 & -0.17869 \\ 0.19443 & -0.69949 & 0.81407 \\ 0.70759 & -0.78886 & -0.01237 \\ 0.87847 & -0.82842 & 0.76975 \end{bmatrix}$$

$$B = \begin{bmatrix} -0.00732 & -0.12021 & -0.17869 \\ 5.11980 & 0.08831 & -0.71657 \\ 7.95398 & 0.00239 & 0.38936 \\ 0.18681 & -0.01173 & 0.02877 \end{bmatrix}$$

From (3) and (4), we can calculate α and β, and substitute them into (2), giving

$$P(d_m < d|N, \sigma) = \frac{1}{1 + \exp[-\frac{(\log d - \alpha(N,\sigma))}{\beta(N,\sigma)}]} \equiv f(d, N, \sigma). \tag{5}$$

Let N_1 denote the solution of $f(d, N_1, \sigma) = P(target)$. For given d, σ, and pre-specified $P(target)$, we approximate N^* by N_1, which is obtained by (5).

To illustrate our approach, we consider σ = 0.004, 0.008,..., 0.04; d = 0.002, 0.008,..., 0.038. The intervals representing the range between the 5th percentile and the 95th percentile for $P(target) = 0.91, 0.94, 0.96, 0.98, 0.992, 0.995, 0.999$ are plotted in Fig 3. It is based on log scale. We set $P(C) = 1 - 10^{-6}$ if there is no disconnection in 10^5 samples.
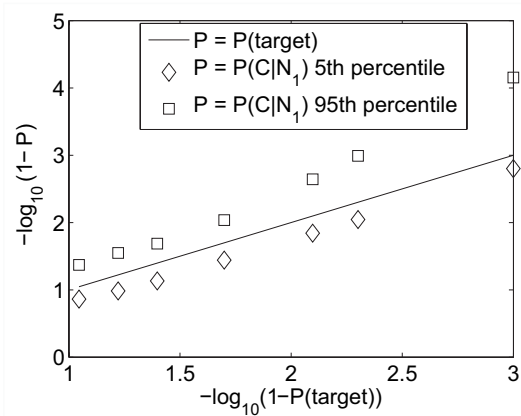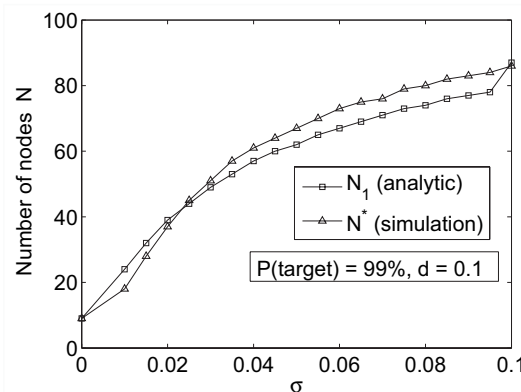
Fig. 3.   Simulation values for $P(C|N_1)$.

Fig. 4.   Analytical and simulation results for N.

Fig. 5.   Behavior of Nd versus d.

Fig. 6. Nd versus d under fixed   .

Fig. 4 compares N_1 calculated from (5) with N^* found from simulations. It shows good agreement. It also shows that even for small σ (small randomness in sensor locations), the value of N^* is much larger than for $\sigma = 0$ (sensors exactly in their optimal locations).
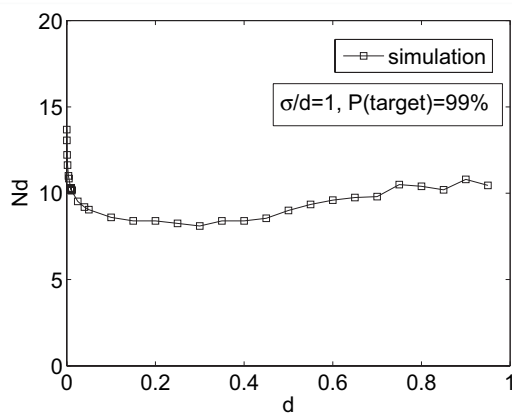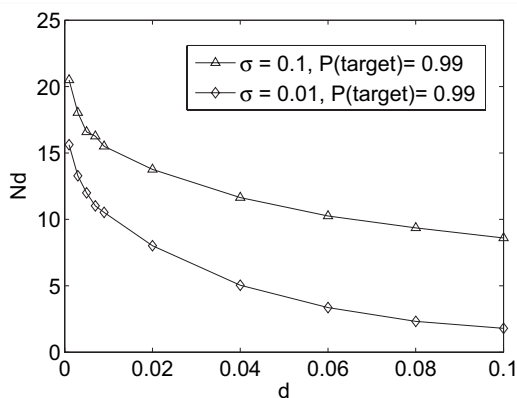
IV. COMPLEMENTARY CASES

Our fitting procedure does not apply if N or σ is small. If N is small, direct simulation is simple. We now consider regimes when σ may be small.

a) Small σ, constant d: The smaller σ is, the more precise are the locations of the sensors. If $\sigma = 0$, $N = D/d - 1$. If $\sigma > 0$, but close to zero, adding a few sensors to the $D/d - 1$ value will provide a good approximation for N.

b) Small d, constant σ/d and $D = 1$: This case is equivalent to the case where d and σ are fixed and D increases to infinity. Fig. 5 shows simulation results for Nd versus d for $\sigma/d = 1$. The value Nd is approximately constant (except for very small d, which correspond to unrealistically many nodes) as indicated by the roughly horizontal line. The intuition is that, given constant d and σ, the number of hops needed, N, is roughly proportional to the distance covered, D. When N is very large, there is an increased probability that one hop is much longer than the mean. This requires an increased margin $d - D/N$, causing Nd to increase if d is very small.

c) Small d and fixed σ: Small d values indicate cheap sensors. It is of interest to study the tradeoff of number of

sensors required versus their cost. The simulation results in Fig. 6 demonstrate that Nd increases as d decreases which indicates that the required number of sensors increases faster than d decreases. If the sensor cost as a function of d is known, results such as those of Fig. 6 could be used to optimize N.

REFERENCES

[1] M. Tubaishat and S. Madria, "Sensor networks: an overview," *IEEE Potentials*, vol. 22, pp. 20–23, April-May 2003.

[2] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 388–404, March 2000.

[3] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing*, 2002, pp. 80–91.

[4] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," in *Proc. 6th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2003, pp. 28–32.

[5] C. Foh and B. Lee, "A closed form network connectivity formula for one-dimensional MANETs," in *Proc. IEEE ICC 2004*, vol. 6, June 2004, pp. 3739–3742.

[6] C. Foh, G. Liu, B. Lee, B. Seet, K. Wong, and C. Fu, "Network connectivity of one-dimensional MANETs with random waypoint movement," *IEEE Commun. Lett.*, vol. 9, pp. 31–33, Jan. 2005.

[7] Y. Cui, K. Xu, J. Wu, and Z. Yu, "An Internet routing emulation system: research and development," in *Proc. International Conference on Communications Technology*, 2003, pp. 495–499.

[8] S. Kotz and L. Norman, Eds., *Encyclopedia of Statistical Sciences*. New York: John Wiley & Sons, 1985, vol. 5.

[9] J. Harris and H. Stocker, *Handbook of Mathematics and Computational Science*. New York: Springer-Verlag, 1998.