99
MW

LA-UR- 00-3340

Approved for public release; distribution is unlimited.

Title: Connectivity of Random Graphs or Mobile Networks: Validation of Monte Carlo Simulation Results

Author(s): L. Jonathan Dowell, TSA-4
Michael Bruno, United States Naval Academy

Submitted to: ACM Symposium on Applied Computing
Las Vegas, NV
11-14 March 2001

Los Alamos
NATIONAL LABORATORY

DISCLAIMER

DISCLAIMER

Connectivity of Random Graphs or Mobile Networks: Validation of Monte Carlo Simulation Results

L. Jonathan Dowell
Los Alamos National Laboratory
P. O. Box 1663 MS F604
Los Alamos, NM 87545
ljdowell@lanl.gov

Michael Lee Bruno
United States Naval Academy
P.O Box 14904
Annapolis, MD 21412
m010735@nadn.navy.mil

ABSTRACT

Distributed sensor networks are systems of sensor nodes interconnected by a communication network. The networks support advanced detection algorithms that use distributed computing to detect events as emergent phenomena of the system of detectors. These systems depend on connected and reliable communication networks. The systems are directly related to mobile networks, with similar characteristics and problems. This paper derives models of these networks based on the theory of random graphs to develop design specifications for node density and communication capability for such networks.

Keywords

distributed sensor network, mobile network, connectivity, random graph

1. INTRODUCTION

Distributed sensor networks (DSNs) are a new type of detection, computation, and communications object. Comprised of an array of discrete sensors, a communications network that links them, and a (possibly distributed) computing environment that combines the measurements of the individual sensors, a DSN supports new types of detection algorithms that can improve detection probability, reduce false alarms, and collect new types of scientific data.

The Anti-Personnel Landmine Alternatives (APLA) Project at the Los Alamos National Laboratory (LANL) seeks to replace certain types of military landmines with an array of sensors across a battlefield, creating a DSN to detect the presence and movement of enemy forces and direct weapons to defeat them. This is an example of a DSN in which sensors are physically distributed at random locations across a geographic area, with the detection and information collection being an emergent property of the sensor system.

The communications network is vital to the operation of the DSN. For the APLA's DSN, the sensor nodes may be linked by radio communications. Then there will be a tradeoff between transmitter power and battery life at the sensor nodes that can be optimized. Important to the optimization of transmitter power to maximize battery life is an understanding of the relationship between transmitter power, communications range, and the connectivity necessary to adjoin the sensor nodes so that signals can be routed efficiently through the communications network.

These problems of DSN optimization are related to recent work in the field of mobile communications networks. As early as 1978, Kleinrock and Silvester [5] developed models for mobile networks that are related to the problems of DSNs. More recently, work has considered message routing by broadcast percolation [1], connectivity as a function of nodal communication distance [6], and optimization criteria for nodal communication distance [9]. Other work in clustering of random nodal phenomena [4] and modeling of mobile networks as a time series of random graphs [8] is also related to this problem.

The purpose of this paper is to explore mathematical models of random communications graphs and to apply those models to the validation of scientific-computing techniques for simulation of DSNs. The analysis will reveal emergent phenomena from the simulation that are predicted by these mathematical models. The work will be used to develop reliability specifications for the design of DSNs through the selection of the number of nodes and their communications capabilities as a function of the area to be covered by the DSN. These reliability specifications will seek to ensure the connectedness of the network, the availability of redundant message-handling paths to support message throughput and tolerate component failures, and the use of simple message-routing algorithms.

2. CONNECTIVITY OF RANDOM GRAPHS

Many communications networks are random graphs. Particularly, distributed sensor networks and mobile networks comprise random graphs through a set of vertices V embedded randomly in a two-dimensional plane (a geographic map) or on a three-dimensional surface (a geographic map with terrain elevation features) such that the edges E connecting these vertices exist if the distance between two vertices is less than some maximum-range parameter and, for the three-dimensional surface, the line-of-sight between the two vertices is unobstructed. Note that a distributed sensor network typically is comprised of a set of fixed vertices, whereas in mobile networks the vertices are able to move to different locations over time. In either case, the set of edges and the resultant network connectivity are emergent properties of the graph that are determined by the locations of the vertices. We will consider the connectivity properties of such random graphs for the purpose of evaluating the ability of the communications network to route messages between pairs of vertices. We will evaluate graphs with fixed vertices, as a model of distributed sensor networks or as a discrete-time "snapshot" of a mobile network.

2.1 Expected Nearest-Neighbor Distance

If the probability density function for the locations of the nodes is known, then other properties of the DSN can be derived. An example is the expected nearest-neighbor distance, the average distance between a node and its closest neighbor.

For example, consider the nodes to be distributed across an area A with uniform probability per unit area α.

$$\int_A \alpha \, dA = 1 \quad \Rightarrow \quad \alpha = \frac{1}{A} \tag{1}$$

Then the probability that the first neighbor of a node is at a distance between r and $r+dr$ is the joint probability of observing a neighbor at this distance and the probability that there are no other nodes closer than the distance r. The expression of this probability uses a combinatorial enumeration of the possible number of nodes in the two regions. This enumeration is a function of the number of nodes, N, that are randomly placed in A to make up the DSN.

$$\begin{aligned}
\Pr\left(1\text{st}\big|_r^{r+dr}\right) &= \Pr\left(0\big|_r\right) \bullet \Pr\left(\text{any}\big|_r^{r+dr}\right) \\
&= \left[1 - \Pr\left(\text{any}\big|_r\right)\right] \bullet \Pr\left(\text{any}\big|_r^{r+dr}\right) \\
&= \left[1 - \sum_{i=1}^{N} \binom{N}{i} \left(\alpha\pi r^2\right)^i \left(1 - \alpha\pi r^2\right)^{N-i}\right] \\
&\qquad \bullet \sum_{i=1}^{N} \binom{N}{i} \int_r^{r+dr} \left(2\alpha\pi x dx\right)^i \int_r^{r+dr} \left(1 - 2\alpha\pi x dx\right)^{N-i} \\
&= \left(1 - \alpha\pi r^2\right)^N \bullet \left[1 - \left(1 - \alpha\pi\left(2rdr + dr^2\right)\right)^N\right]
\end{aligned} \tag{2}$$

Using a binomial expansion of the term involving dr,

$$\begin{aligned}
\Pr\left(1st\big|_r^{r+dr}\right) &= \left(1 - \alpha\pi r^2\right)^N \left[1 - \left(1 - 2N\alpha\pi(rdr + dr^2) + H.O.T.\right)\right] \\
&= \left(1 - \alpha\pi r^2\right)^N \left[2N\alpha\pi rdr + H.O.T.\right]
\end{aligned}, \tag{3}$$

where $H.O.T.$ signify higher-order terms of powers of dr. Then the probability density function for the nearest-neighbor distance, $P(r)$, is

$$P(r) = \lim_{dr \to 0} \frac{\Pr\left(1st\Big|_{r}^{r+dr}\right)}{dr} \qquad (4)$$
$$= 2N\alpha\pi r\left(1 - \alpha\pi r^2\right)^N$$

The expected nearest-neighbor distance, $\langle r \rangle$, is a useful metric for establishing a lower bound for the necessary communication range, so that a node in the DSN will at least be able to communicate with its nearest neighbor. It also is a performance metric for the evaluation of DSNs through simulation.

$$\langle r \rangle = \int_0^{R(A)} rP(r)dr = \int_0^{R(A)} 2N\alpha\pi r^2\left(1 - \alpha\pi r^2\right)^N dr$$
$$= \frac{-r\left(1 - \alpha\pi r^2\right)}{2\alpha\pi(N+1)}\Bigg|_0^{R(A)} + \int_0^{R(A)} \frac{\left(1 - \alpha\pi r^2\right)^{N+1}}{2\alpha\pi(N+1)} dr$$
$$= \frac{1}{2\alpha\pi(N+1)} \sum_{k=0}^{N+1} \binom{N+1}{k} \frac{\left(-\alpha\pi r^2\right)^k r}{2k+1}\Bigg|_0^{R(A)} \qquad (5)$$
$$= \frac{A^{\frac{1}{2}}}{2\alpha\pi^{\frac{3}{2}}(N+1)} \sum_{k=0}^{N+1} \frac{(-1)^k}{2k+1}$$

2.2 Expected Number of Neighbors

A similar analysis determines the expected number of neighbors within a nodes communications range, R. The node is able to communicate over an area πR^2, so the probability that any one other node falls within this area is $\alpha\pi R^2$. The probability that there will be k nodes within this area out of N-1 possibilities is

$$\Pr\left(k\big|_{N-1}\right) = \binom{N-1}{k}\left(\alpha\pi R^2\right)^k\left(1 - \alpha\pi R^2\right)^{N-1-k} \qquad (6)$$

Then the expected number of neighbors, $\mu = \langle k \rangle$, is

$$\mu = \sum_{k=0}^{N-1} k \Pr(k) = \sum_{k=0}^{N-1} k \binom{N-1}{k} (\alpha\pi R^2)^k (1-\alpha\pi R^2)^{N-1-k} = (N-1)\alpha\pi R^2 \quad (7)$$

2.3 Holes in Communications-Network Coverage

It is useful to consider the probability that there is an area B within the DSN that contains no nodes. This probability is

$$\Pr(0,B) = (1-\alpha B)^N \quad (8)$$

Connectivity through the network depends on adequate communication between neighboring nodes to repeat and route messages through the network. If there is an area containing no nodes and the extent of that area exceeds the communication distance of the neighboring nodes, then no communication will occur across that area. The area will constitute a "hole" in the network, requiring some messages to be routed longer distances to bypass the hole or possibly disconnecting the network by dividing it into two unconnected subnetworks.

3. SIMULATION OF RANDOM GRAPHS

Some analyses of DSNs can be performed through computer simulation. Particularly, Monte Carlo techniques can be used to simulate the random locations of nodes in a DSN for the purpose of evaluating network connectivity and other aspects of the network as a graph.

3.1 Graph Visualization Tool

Figures 1, 2, and 3 illustrate a DSN graph visualization tool that has been developed by the APLA as a Java™ applet for the world-wide web. This applet is presently available on the LANL website, at http://public.lanl.gov/u106527/DSN/GrphDemo/GrphDemo.html. These figures demonstrate the utility of this tool in evaluating the communication network as a function of the nodal communication distance, R. Clearly, if R is too small, then the network will not be connected. Although connectivity and network message routing are improved with increasing R, the energy requirement, battery cost, and expected battery life of these nodes place a premium reducing R. Therefore, there will be an optimum communication distance for a particular DSN, and this optimum value can be evaluated by computer simulation.

3.2 Neighbor-Location Probability Density Function

The results developed in Section 2 can be used as validation of the simulation results. Figure 4 compares the empirical nearest-neighbor probability density function observed in the simulation with the expression derived in Eq. 4. The simulation evaluated the random location of 400 nodes, repeating the procedure through 100 trials. The nearest-neighbor distance for each node was recorded, and statistics were calculated for the fraction (that is, the empirical probability) of observing a nearest-neighbor distance between r and $r+\Delta r$. These results are compared with the probability density function by integrating that function to determine the probability of observing a nearest neighbor within that distance interval. Figure 4 shows the excellent agreement between the simulation and the calculation.

3.3 Connectedness of Random Graphs

A crucial objective of the communications-network aspect of the DSN is that the network be connected. Through simulation, we evaluated a number of networks, varying both the number of sensors and the nodal maximum communication radius to evaluate the dependence of network connectivity on these design parameters. Table 1 shows results of these simulations. As expected, the probability that all nodes in the DSN were connected increased as the nodal maximum communication radius increased. Below a finite radius threshold, few of the networks were completely connected. However, above this threshold nearly all the networks were connected. This radius threshold was larger for sparse DSNs containing fewer nodes in an area.

Figure 5 illustrates a model for computing the probability of non-trivial (e.g., more than one node disconnected). In this model, a node is capable of communicating with neighbors if they are within the node's communications radius. For simplicity of tiling, let this radius be approximated by a hexagon, as shown in the figure. Then a "hole" in the DSN is an area containing no nodes such that no communications edges (drawn as straight lines between adjacent neighbors) cross this area. There are several ways such holes can be described, but for simplicity let us partition the DSN's area of coverage by a pattern of regular hexagons of radius (and side length) equal to the nodes' communications radius. With this approximation, a simple description for the probability of a hole (comprised of a set of unoccupied hexagons) capable of dividing the network into to disconnected subgraphs.

The probability that one hexagon is unoccupied by nodes is

$$\Pr(0) = (1 - \alpha A_H)^N = \left(1 - 3\sqrt{3}\alpha R^2\right)^N, \qquad (9)$$

where A_H is the area of the hexagon. Then the joint probability that six such unoccupied hexagons are arranged in the pattern shown in Fig. 5 (the minimum number of unoccupied hexagons away from the edges of the DSN area capable of causing a disconnected network) is

$$\Pr\begin{pmatrix}\text{six disconneting} \\ \text{hexagons}\end{pmatrix} = \left[\left(1 - 3\sqrt{3}\alpha R^2\right)^N\right]^6. \qquad (10)$$

Note that this joint probability becomes vanishingly small with R and N sufficiently large. (Although there are many possible locations for such a set of disconnecting holes in the DSN, the probability of the occurrence of each of them is the same and will be a function of the expression given in Eq. 10. This expression is a useful metric for considering the possibility of disconnection of a particular DSN.) Thus, for sufficient R and N, nearly all DSNs with those parameters will be expected to be connected, supporting the results described in Table 1.

The presence of holes in the DSN area also can necessitate indirect shortest paths for routing messages between pairs of nodes. Figure 6 shows how two holes increase the path length and number of retransmissions necessary for communication between nodes C and D.

4. FURTHER WORK

Additional work is necessary to develop the model of holes in the DSN area. This work includes the development of techniques for evaluating holes of irregular shape, combinatorics of the number of possible hole configurations for disconnecting a network, and models for the expected path length and number of retransmissions between pairs of nodes in the DSN.

5. ACKNOWLEDGMENTS

Los Alamos National Laboratory (LANL) supported this work through the APLA project. Michael Bruno was supported by the LANL Service Academy Research Associates program under sponsorship by the United States Naval Academy. The University of California operates LANL for the United States Department of Energy under contract *W-7405-ENG-36*. LANL strongly supports academic freedom and a researcher's right to publish; therefore, LANL as an institution does not endorse the viewpoint of a publication or guarantee its technical correctness.

Sun® and Java™ are trademarks of Sun Microsystems, Inc.

6. REFERENCES

[1] Y. Cheng and T. G. Robertazzi, Critical Connectivity Phenomena in Multihop Radio Models, IEEE Trans. Commun. 37 (1989) 770-777.

[2] L. J. Dowell, Optimal Configuration of a Command and Control Network: Balancing Performance and Reconfiguration Constraints, Proc. 2000 ACM Symp. Appl. Comp. (Como, Italy, Mar. 2000), ACM, 94-98.

[3] P. Gupta and P. R. Kumar, The Capacity of Wireless Networks, IEEE Trans. Info. Theory 46 (2000) 388-404.

[4] A. M. Kellerer, On the Number of Clumps Resulting from the Overlap of Randomly Placed Figures in a Plane, J. Appl. Prob. 20 (1983) 126-135.

[5] L. Kleinrock and J. A. Silvester, Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number, Conf. Rec. Nat. Telecommun. Conf. (1978) 4.3.1-4.3.5.

[6] J. Ni and S. A. G. Chandler, Connectivity Properties of a Random Radio Network, IEE Proc. Commun. 141 (1994) 289-296.

[7] Z. Shen, The Calculation of Average Distance in Mesh Structures, Proc. 2000 ACM Symp. Appl. Comp. (Como, Italy, Mar. 2000), ACM, 89-93.

[8] P. D. Stroud, Kalman Extension of the Genetic Algorithm, Proc. SPIE 4055 (2000) 176-187.

[9] M. W. Subbarao and B. L. Hughes, Optimal Transmission Ranges and Code Rates for Frequency-Hop Packet Radio Networks, IEEE Trans. Commun. 48 (2000) 670-678.

[10] S. M. Sussman, A Survivable Network of Ground Relays for Tactical Data Communications, IEEE Trans. Commun. 28 (1980) 1616-1624.

7. ABOUT THE AUTHORS

Dr. L. Jonathan Dowell is a Technical Staff Member in the Energy and Environmental Analysis Group at LANL. Dr. Dowell received his Ph.D. in engineering physics from the University of Virginia in 1989. Dr. Dowell has performed research and analysis in the fields of metrology, power-systems engineering, computer science, and applied mathematics. Dr. Dowell also is president of ReefNews, Inc., a non-profit organization that produces educational materials about the oceans.

Michael Bruno is a 1/C Midshipman at the United States Naval Academy. Midn. Bruno is a mathematician and is Captain of the USNA 2000-2001 Debate Team.

There is a relationship between **Node Density** and **Comm. Range** for a graph to be connected with high probability.

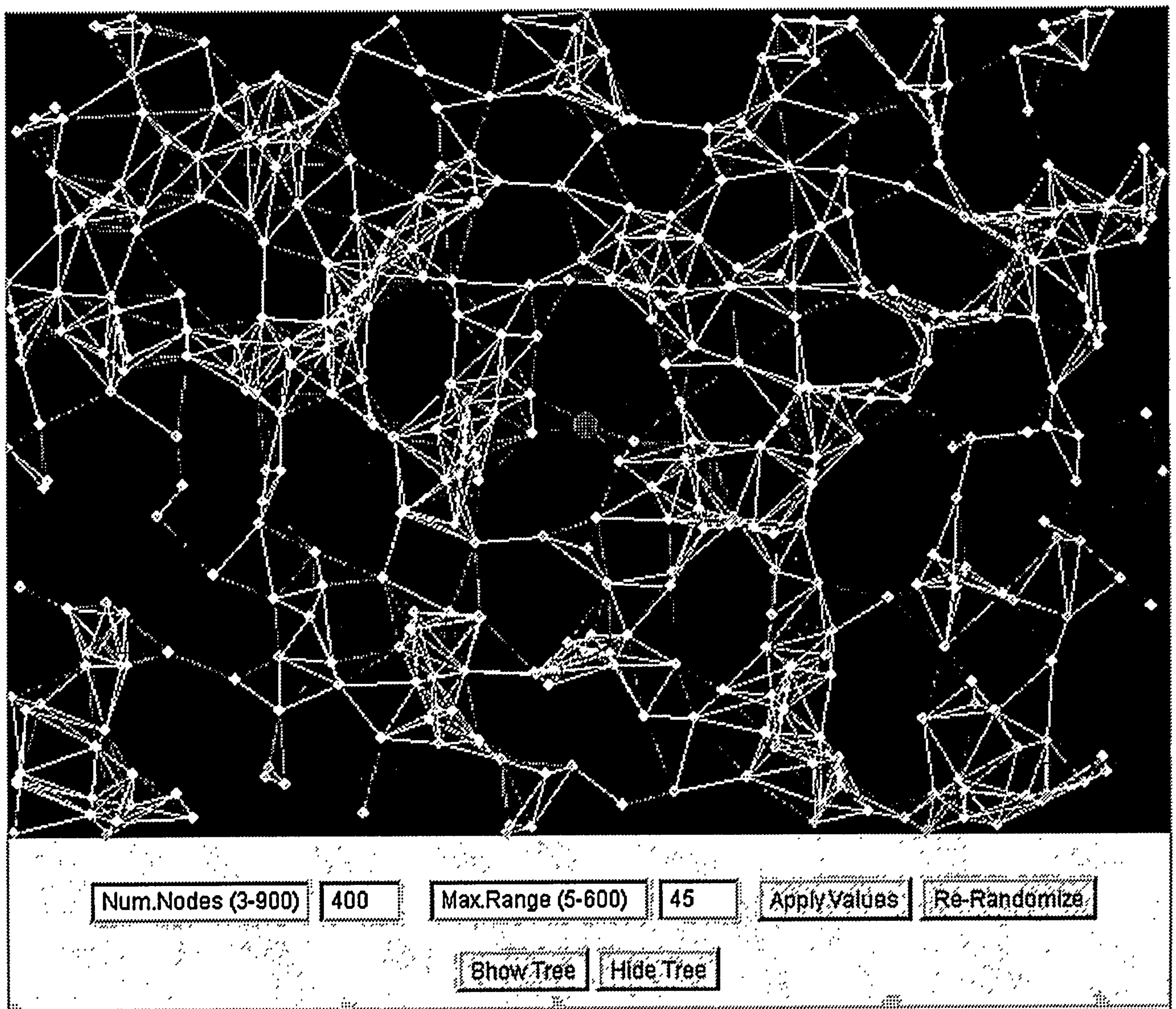

Figure 1

A graph with **Good Node Density** but **Poor Comm. Range** may be disconnected.

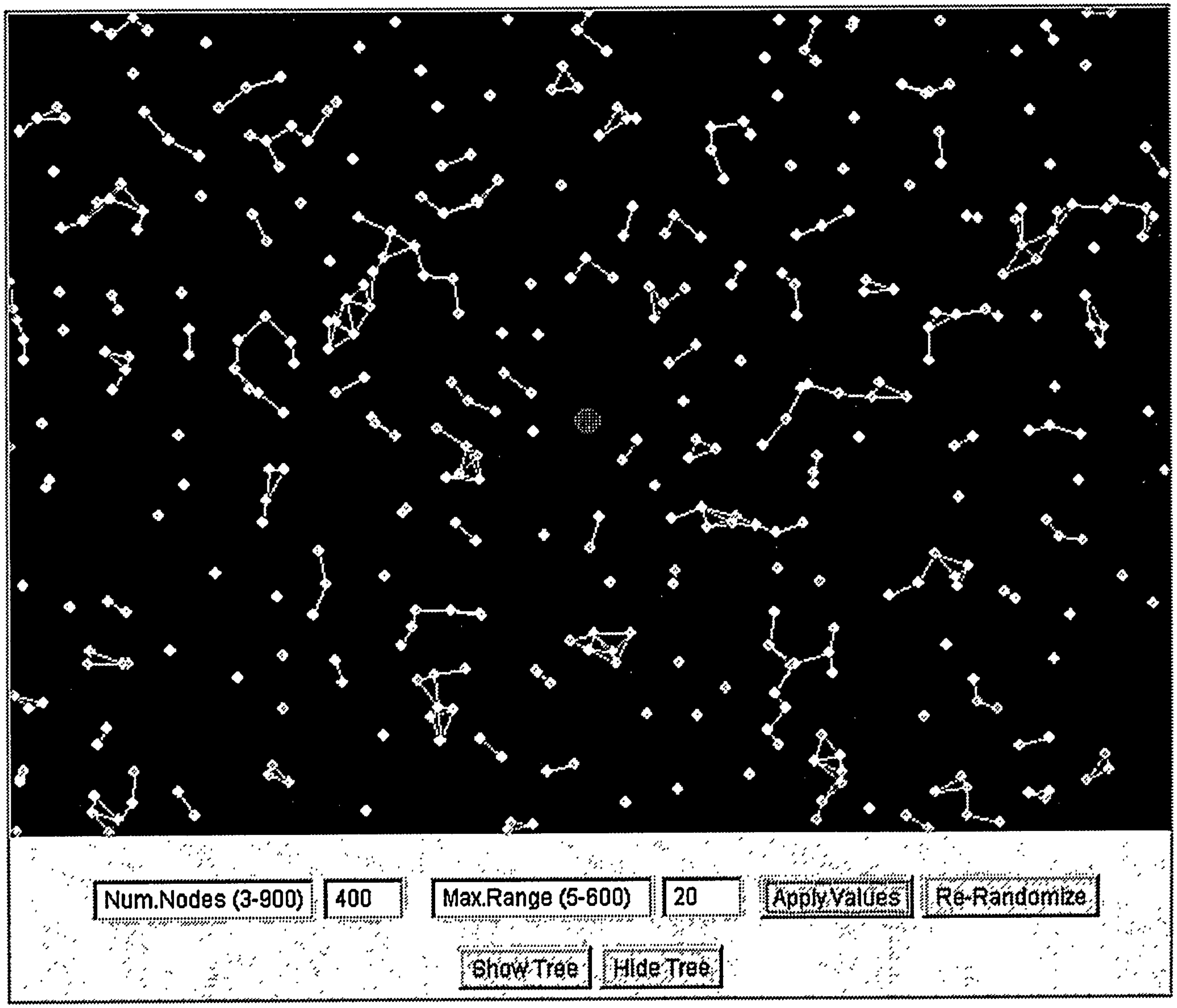

Figure 2

Increasing communications range can degrade system performance from computational burden of large number of options.

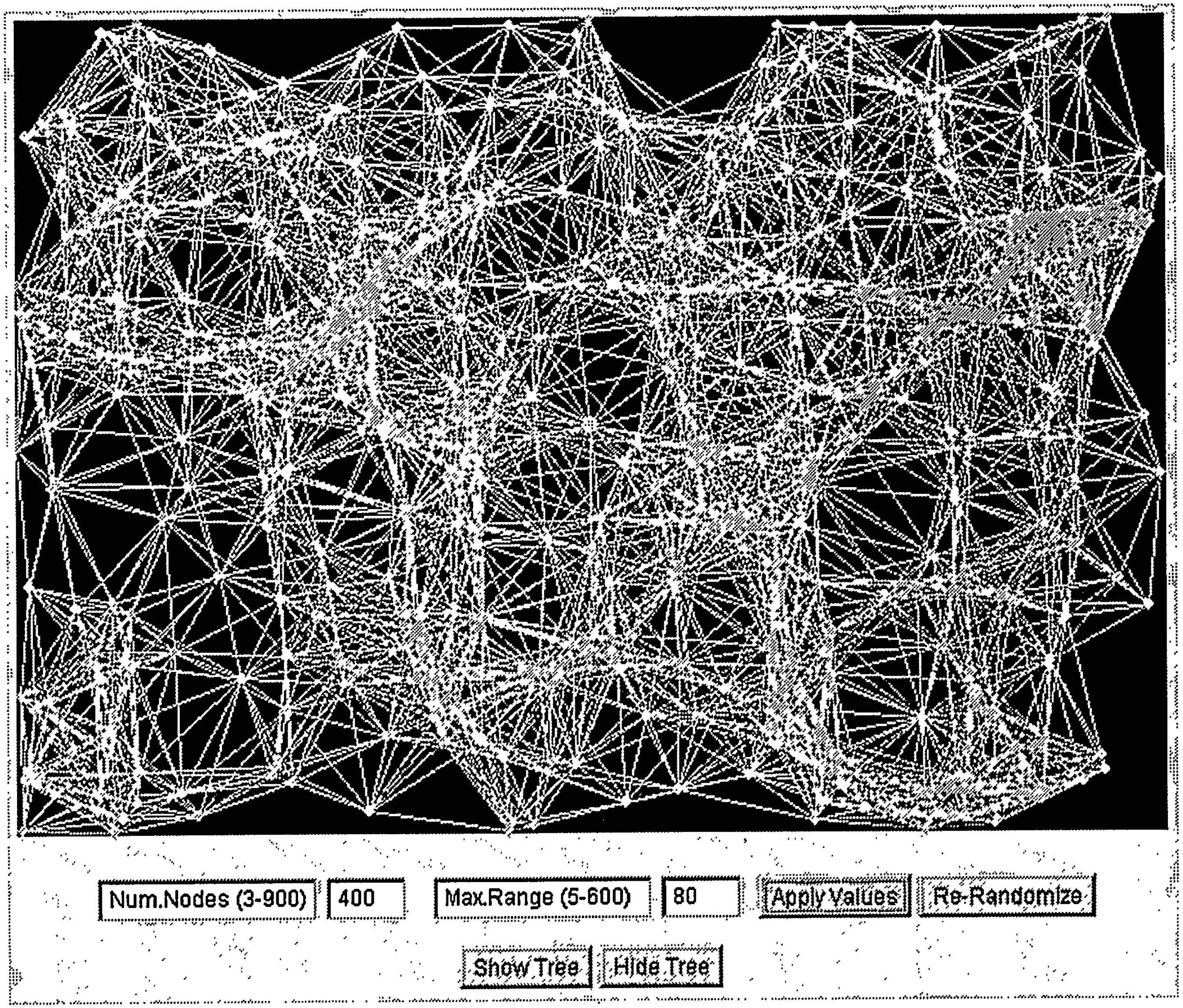

Figure 3

The Distribution-Function is an Accurate Model of the Simulation Result.

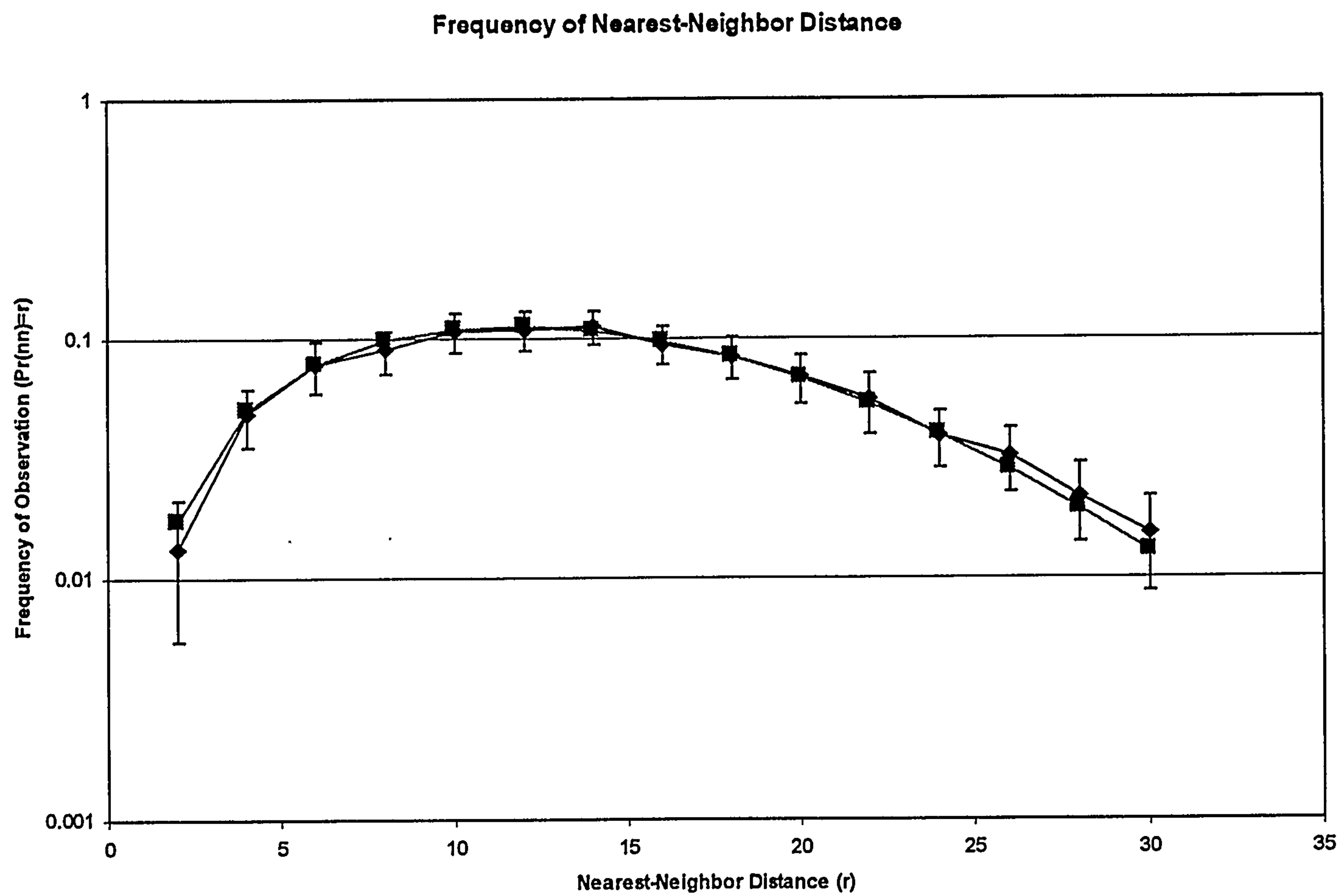

Figure 4

Simulations reveal the relationship between **number of nodes**, **comm. range**, and **graph connectivity**.

Number of Complete Graphs (out of 200 tries)					
	Range				
Nodes	10	20	50	100	200
10	0	0	0	0	13
20	0	0	0	0	40
50	0	0	0	0	191
100	0	0	0	30	199
200	0	0	0	182	200
500	0	0	56	200	200
Average Number of Nodes Connected to CP					
	10	20	50	100	200
10	1.01	1.02	1.20	1.73	5.64
20	1.00	1.05	1.34	3.25	16.53
50	1.05	1.10	1.85	18.82	49.92
100	1.09	1.24	4.45	91.81	99.99
200	1.13	1.68	31.64	199.80	200.00
500	1.32	3.56	496.26	500.00	500.00
Average Fraction of Graph Connected to CP					
	10	20	50	100	200
10	0.10	0.10	0.12	0.17	0.56
20	0.05	0.05	0.07	0.16	0.83
50	0.02	0.02	0.04	0.38	1.00
100	0.01	0.01	0.04	0.92	1.00
200	0.01	0.01	0.16	1.00	1.00
500	0.00	0.01	0.99	1.00	1.00
Average Radius (number of hops back to CP)					
	10	20	50	100	200
10	0.00	0.01	0.10	0.31	1.19
20	0.00	0.02	0.16	0.75	1.96
50	0.03	0.05	0.34	2.80	2.00
100	0.04	0.12	1.00	4.65	1.94
200	0.07	0.28	4.16	3.85	1.90
500	0.15	0.83	8.40	3.48	1.88

Table 1

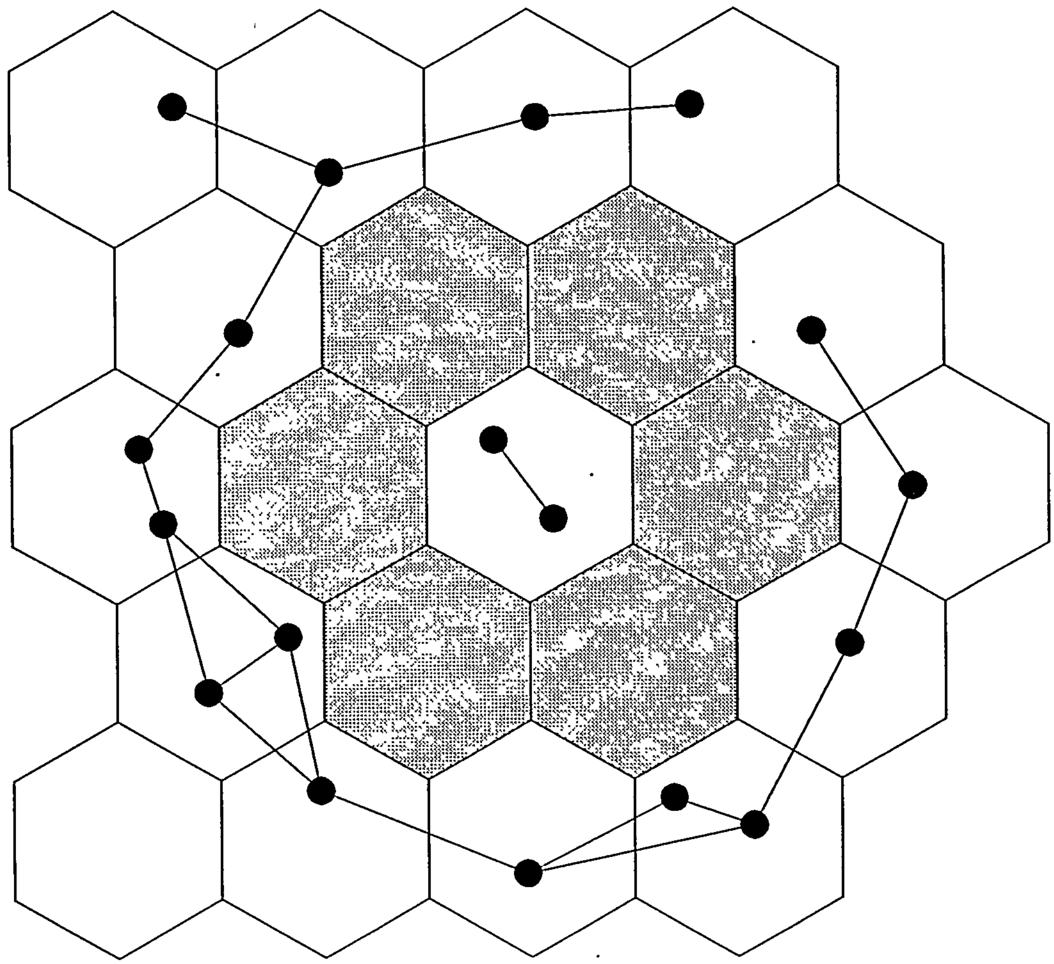

Figure 5. The presence of "holes" in a DSN (indicated by shaded hexagons) is a necessary condition for the network to be disconnected.

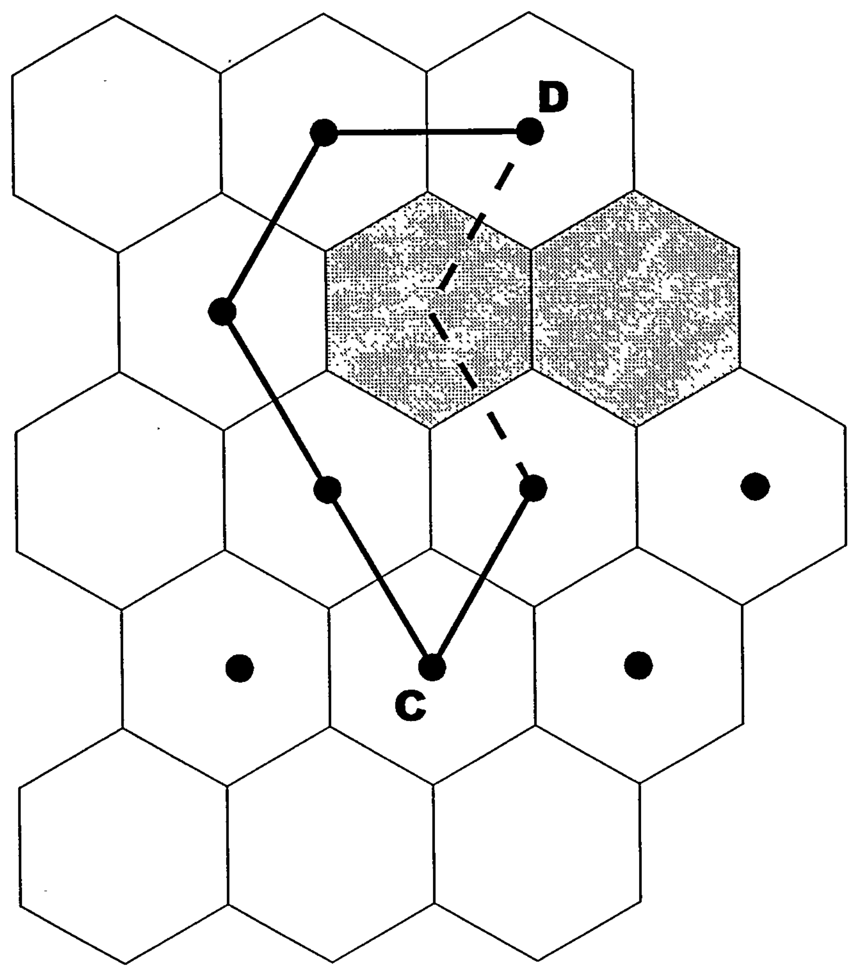

Figure 6. Holes also can lengthen message routing paths and necessitate complicated routing algorithms.