Abstract—Networked systems are essential to the function of modern society and the consequences of damage to networks can be severe. Assessing the performance of a network is an important step for recovering damaged networks and designing reliable networks. Some of the key general indicators of network performance are connectivity, distance between node pairs, and number of alternative routes. We focus on sensor networks with a topology modeled by a class of random geometric graphs (RGGs). In order to evaluate survivability and reliability, we consider two types of failure modes in a RGG: uniform and localized node failures. Since network performance is multi-faceted, and assessment can be time constrained, we introduce four measures, each of which can be computed in polynomial time, to estimate performance of a damaged RGG. Theoretical analysis of these four measures is challenging, especially when the underlying graph becomes disconnected. The focus of this paper is to conduct simulation experiments on several measures of network performance through the temporal process of node failures. Together with the empirical results the performance measures are analyzed and compared in order to provide understanding of the two different failure scenarios in a RGG.

I. INTRODUCTION

Advances in sensor technology have enabled a wide range of wireless sensor network applications with sensors distributed in the field using decentralized control and communication algorithms. These applications include battlefield surveillance, physical structure monitoring, and environmental monitoring [1]–[3]. Analyzing the reliability of communication in these networks has become an active research area, and measuring and designing performance is a major goal.

In sensor networks, edges represent the communication channels, or connection, among nodes. Loss of connections can be due to factors such as interference, weather, or field obstacles. The sensors themselves can fail due to battery depletion, natural disasters, harsh environmental conditions, or even targeted attacks. Similar conditions can interfere and even disrupt the transmissions among sensors. It is important to understand the performance of a sensor network under such failures.

The focus of this work is to investigate measures that capture both local and global performance characteristics of wireless sensor networks under different node failure scenarios. Our approach is to use a random geometric graph (RGG) model in which nodes (and the incident edges) fail randomly according to a specified probability distribution. RGGs have been a standard tool to model and study wireless ad-hoc and wireless sensor networks [4], [5] since they capture the generic topological features of communication. In order to study dynamical node failures, simpler models, such as trees, have been used [6]. Even in simplified models it is challenging to derive theoretical results for measuring network performance under node failures.

The main contribution of this paper is to use simulation to investigate network failures under less than ideal situations. We use four existing network measures to evaluate performance and show which of those capture the effects of uniform vs. localized node failures. Though the networks we use are only simplified models of real-world sensor networks, these results can help to guide analysis and measurement of failures in more realistic systems and systems where detailed knowledge of the network is not available. In Sec. II we review some of the previous results on network reliability and resilience. We describe the model and failure scenarios in Sec. III. The set of four performance measures is described in Sec. IV, and in Sec. V we evaluate and analyze those measures through simulations of the different stages in a failing network and under different failure modes.

II. RELATED WORK

A. Reliability and Resilience

Determining the reliability of networks is vital for planning communication systems, and there has been extensive research in wired and wireless systems [7]–[10]. The well-studied network reliability problem considers a probabilistic graph where each edge and/or node can fail independently with a given probability, and the goal is to calculate the probability that communication can be established among the set of surviving nodes. The analysis of network reliability has guided the design of telecommunication networks, e.g., introducing redundant links to guarantee reliability [11]–[13].

In the two-terminal problem [14]–[16], with two special nodes called source s and destination t, the goal is to find the probability that there exists a path between these two nodes. AboElFotoh et al proved the two-terminal reliability problem is #P-complete for RGGs and grids, and showed methods for computing upper and lower bound of the reliability [17]. More
general cases involve k-terminal and all-terminal reliability problems \([11], [18]–[20]\). The network reliability problem was proved to be \#P-complete \([11], [21], [22]\), but Monte Carlo simulation and bounding methods have been demonstrated to be two efficient ways to solve this problem \([9], [13]\).

In \([23], [24]\), Colbourn introduced network resilience, which is the expected number of node pairs that can communicate under independent edge failure. Computing the network resilience is \#P-complete for a general and planar networks. The resilience of \(n\)-node series-parallel networks can be computed in \(O(n^2)\) time \([23]\).

### B. Random Geometric Graphs with Node or Edge Failures

The RGG model with node or edge failures has been studied by Díaz et al. in \([25]\). Several network measures, Bisection, Minimum Linear Arrangement, and Minimum Cutwidth, were considered. The main results show that RGGs could tolerate a constant edge or node failure probability while preserving the order of magnitude of the measures, and that there is a Hamiltonian cycle asymptotically with probability one, given a constant (node or edge) failure probability.

Kong and Yeh studied a model of wireless networks where the failure of a small number of nodes can cause global failure effects \([26]\). In that model the node failure probability depends on node degree and the cascading failure problem becomes equivalent to a degree-dependent site percolation process on a RGG.

Asymptotic results for the existence of the giant component (the largest connected component containing a positive fraction of all nodes) and complete connectivity of RGGs with \(n \rightarrow +\infty\) nodes and radius \(r = r(n)\) as a function in \(n\) are given in \([5], [27]\). The transition from fully connected RGGs to fully partitioned graphs under uniform node failures has been analyzed in \([28]\), and two measures were introduced: the last connection time (the last time that the network keeps a majority of surviving nodes connected in a single giant component) and the first partition time (the first time that the remaining surviving nodes are partitioned into multiple small components). Xin and Wan have shown that these two measures are of the same order \([28]\).

### III. Modeling Node Failures

#### A. Random Geometric Graph Model

A wireless sensor network typically contains a large number of randomly (in space) deployed nodes with links determined by geometric proximity (radio range) among the nodes. That is, a communication link between two nodes exists if the geometric distance between them is sufficiently small to enable successful signal transmission. The RGG model captures this structure \([4], [5]\); nodes are randomly placed in \(\mathbb{R}^d\) and connected if they are within a threshold distance \(r\) of each other. The geometric nature of RGGs allows us to study node failure scenarios that are uniform and nonuniform in space.

**Definition 1.** (Random Geometric Graph Model \([27]\)) For the \(d\)-dimensional space \(\mathbb{R}^d\) provided with the distance norm \(| \cdot |\), let \(\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\} \subset [0,1]^d\) be chosen independently and uniformly at random. The random geometric graph \(G(\mathcal{X}_n; r)\) has the node set \(\mathcal{X}_n\) and the edge set \(E = \{(X_i, X_j) : i \neq j, X_i, X_j \in \mathcal{X}_n, |X_i - X_j| \leq r\}\).

In the following we study the \(d = 2\) dimensional case.

#### B. Node Failure Scenarios

To investigate how network performance changes under node failures we consider two failure scenarios: uniform and localized (centralized). These scenarios are of practical interest in wireless ad-hoc networks, such as a network of autonomous sensors that cooperatively monitors physical or environmental conditions, or a set of PDA (personal digital assistant) devices that cooperatively perform some computation. Spatially localized failures, such as those due to a natural disaster or a targeted attack, are modeled as centralized failures, and natural device malfunctions are modeled with uniform failures.

For a given scenario nodes are each assigned a specified failure probability. Network failure is simulated by stepping in time with each node failing independently according to the failure probability at each time. When a node fails the node and all incident edges are removed from the current network. We formally define the two failure scenarios as follows.

**Definition 2.** (Uniform Scenario) The Uniform Random Node Failure is an i.i.d. process, where any node \(i\) in a graph fails with the same probability \(p_i = p\), where \(p \in [0, 1]\).

**Definition 3.** (Localized Scenario) The Spatially Localized Random Node Failure is an i.i.d. process, where any node \(i\) in a graph fails with the probability \(p_n\), which is obtained from Gaussian distribution \(\lambda \exp(-d_i^2/(2\sigma^2))\). Here \(d_i = \|x_i - (1/2, 1/2)\|\) is the Euclidean distance of the position \(x_i\) of the node \(i\) from the center \((1/2, 1/2)\), and \(\lambda \leq 1\) is a positive constant.

In the Uniform Scenario the residual network at each time step is still a RGG since nodes are failed uniformly at random. But in the Localized Scenario after any node removal the remaining network will not likely be a RGG. Our paper focuses on comparing these two failure scenarios over a set of network performances measures.

### IV. Performance Measures

To measure network performance under the failure scenarios we consider several measures that incorporate local and global properties related to connectivity and paths in the network.

#### A. Number of Components

If the network is not connected communication between all pairs is not possible. We track the number of connected components over time \(K(t)\) as the nodes fail by using a simple breadth-first search algorithm.

In the Uniform Scenario the nodes fail uniformly at random over time so the lifetime of a node is a geometric random variable \(\text{Geom}(p)\). Thus the expected lifetime of any node is \(E[T] = 1/p\). At the beginning \(r\) is chosen such that the
Fig. 1. A RGG with \( n \) nodes is connected with probability one as \( n \to +\infty \). The original graph is connected with probability one as \( n \to +\infty \). The nodes fail independently and uniformly in space resulting in a sequence of RGGs in time, denoted as \( G_t \), with \( n(t) \) nodes and with the same given radius \( r \). We now estimate the time step \( t \) when the graph \( G_t \) becomes disconnected. A RGG \( G_t \) is disconnected, with the probability one as \( n(t) \to +\infty \), when \( n(t) \) and \( r \) satisfy \( n(t) \pi_r^2 = \log n - o(1) \) [5], [27]. In the Uniform Scenario \( n(t) \) is decreasing in \( t \). Thus, for \( t \) such that \( n(t) < x \), where \( x \) is the unique root of the equation \( \log x/x = \pi_r^2 \) in \( [e, +\infty) \), the current graph \( G_t \) becomes disconnected with probability one as \( n \to +\infty \).

In the Localized Scenario instead of \( p_i = p \) we have \( p_i = \lambda \exp(-d_i^2/2\sigma^2) \). The lifetime of a node \( i \) is still a geometric random variable \( \text{Geom}(\pi_i) \) with the expected value \( \mathbb{E}[T_i] = 1/\pi_i \). In this case the nodes fail independently, but not uniformly, in space so the graphs obtained after failures are not RGGs and the time when the graph becomes disconnected must be estimated numerically.

B. Average Shortest Path

Wireless sensor networks generally operate under constrained energy environment, and the shortest path between two nodes indicates the required energy for conveying data [29], [30]. The average of all-pairs shortest paths provides a global measure of the best connectivity across the whole network. At time \( t \) let \( S(t) \) be a set of all pairs of connected nodes. For a pair of nodes, \((u,v) \in S(t)\), the length of the shortest path \( d(u,v) \) can be calculated since \( u \) and \( v \) are in a connected component. The simple average shortest path measure is taken over all pairs of connected nodes,

\[
H(t) = \frac{1}{|S(t)|} \sum_{(u,v) \in S(t)} d(u,v). \tag{1}
\]

The value of \( H(t) \) measures how well nodes communicate pairwise in a connected component, but discounts for the size of a component. For example a connected graph with 100 nodes that does not form a complete graph has a higher \( H(t) \) than 50 graphs of size 2.

For a connected \( G_t \), the expected value \( \mathbb{E}[H(t)] \) of the average shortest path can be estimated as follows. For any fixed node \( u \) and any node \( v \) disregarding border effects, and given the distance \( \rho := ||X_u - X_v|| \), we have \( d(u,v) = \Theta(\rho/r) \) [31]. Therefore,

\[
\mathbb{E}[H(t)] = \frac{1}{|S(t)|} \mathbb{E}[\Theta(\rho/r)] = \frac{1}{|S(t)|} \sum_{u \in V} \mathbb{E}[\Theta(\rho/r)] = n(t)(n(t) - 1) \cdot \Theta \left( \int_0^{1/\pi} \frac{1}{\pi} \frac{1}{r} \rho^2 2\pi \rho d\rho \right) = \Theta(1/r), \tag{2}
\]

where we used \( |S(t)| = (n(t)/2) \) since every pair of nodes is connected.

C. Component-Averaged Shortest Path

To reflect the fact that the network may become disconnected after failures we introduce a measure which combines the sizes of connected components with shortest paths. At each time \( t \) we compute the connected components \( C_1(t), \ldots, C_{K(t)}(t) \), where \( K(t) \) is the number of connected components at time \( t \). We then calculate a normalized Wiener index [32], that is, we compute the mean value of all-pairs shortest paths \( M_i(t) \) for each component \( C_i(t) \) individually,

\[
M_i(t) = \frac{1}{|C_i(t)|} \sum_{u,v \in C_i(t)} \frac{d(u,v)}{|C_i(t)|^2}. \tag{3}
\]

Finally we introduce the value \( W(t) \) that combines both the Wiener index and the size of each component,

\[
W(t) = \sum_{i=1}^{K(t)} \frac{M_i(t)}{|C_i(t)|}. \tag{4}
\]

The measure \( W(t) \) penalizes effects of breaking a network into smaller components. For a given Wiener index \( M_i(t) \) in a component a larger component size \( C_i(t) \) results in smaller ratio \( M_i(t)/|C_i(t)| \) and hence better (lower) \( W(t) \). We can also see that the value \( M_i(t)/|C_i(t)| \) reflects the fraction of lost node pairs after breaking a network into components.
D. Communicability

Network communicability [33] was designed to reflect the fact that communication does not always occur on the shortest pathways in a network. Communicability is a special case of the Katz measure [34] that computes the number of walks (possibly with loops) through powers of the network adjacency matrix. The shortest-path based measures can be viewed as a special case of the communicability where only a shortest path has positive weight and all others have weight zero. The definition of communicability specifies factorial weights, giving larger weight to the shorter walks and smaller weight to the longer walks [33]. Formally, if \( A(t) \) is the adjacency matrix of a graph \( G_t \), then the communicability \( Z(t) \) is defined as

\[
Z(t) = e^{A(t)} = \sum_{\ell=0}^{\infty} \frac{A(t)^{\ell}}{\ell!}.
\]

Notice that for each pair of nodes \((i,j)\) the entry \( Z(t)_{ij} \) represents the weighted sum of all walks, possibly with loops, starting at \( i \) and ending at \( j \). In the following experiments we plot the logarithm value of the sum of all the entries in \( Z(t) \),

\[
Q(t) = \log \sum_{0 \leq i,j \leq n(t)} Z(t)_{ij}.
\]

Higher communicability \( Z(t) \) indicates better average performance of the network.

V. EXPERIMENTAL RESULTS

In this section we present simulation results. It is challenging to measure the deteriorating performance in a network with node failures and is unlikely that a single network measure can be used to capture all of the differences in the failure scenarios. In our experiments the four measures discussed in the last section are used. The results show that these four measures can compensate each other and provide some basis for comparisons.

A. Setting and Description

In our computational experiments, a RGG \( G(\mathcal{P}_{1000};0.06) \) of 1000 nodes and the radius \( r = 0.06 \) is generated as an underlying network (see Fig. 1). The node failure probabilities are calculated according to a specified failure scenario as described in Sec. III-B. Each simulation is performed over 300 discrete time steps. At time step \( t \), nodes and incident edges are removed from the graph independently with the specified node failure probabilities. The removal of nodes and edges results in a damaged network \( G_t \) at time step \( t \). For \( G_t \) a set of measures including the number of connected components \( K(t) \), average shortest path \( H(t) \), component-averaged shortest path \( W(t) \) and total network communicability \( Q(t) \) are calculated to characterize the damage.

We analyze the failure performance with respect to the number of nodes failed, \( F(t) \), rather than time. That allows us to account for various values of \( p \) in the uniform failure scenario, or \( \lambda \) in the localized failure scenario, together since they only result in a rescaling of time. For example given \( \hat{p} = \alpha p \), the probability that a node fails within \( \hat{t} \) steps with probability \( p \) is \( 1 - (1 - p)^\hat{t} \) and within \( \hat{t} \) steps with probability \( \hat{p} \) is \( 1 - (1 - \alpha p)^\hat{t} \). A simple calculation shows that scaling \( \hat{t} = t \log(1 - p)/\log(1 - \alpha p) \) gives the same expected node failure curves. Because of this rescaling we see that by plotting the performance measures versus the number of nodes failed we only need to consider one value of \( p \) and that the parameter value \( \lambda \) is not significant (other than to set the timescale of failures). In the localized failure scenario we still must study various values of the parameter \( \sigma \) since we cannot rescale space in the same simple way.

B. Case Analysis

Five failure scenarios are considered in the following case analysis. The scenarios include uniform failure \( p = 0.005 \), and localized failures with the variance \( \sigma = 0.15, 0.25, 0.5, 0.75 \). Lower variance \( \sigma \) indicates a smaller (more localized) high probability failure area and larger \( \sigma \) produces a more spatially uniform failure distribution. For each scenario we average over an ensemble of 100 simulations. The resulting network measures are plotted with respect to the number of failed nodes \( F(t) \) in Fig. 2.

![Network performance for failures in a RGG (n = 1000, r = 0.06) under 5 different node failure scenarios. The network performance measures are shown versus the number of failed nodes F(t) and averaged over an ensemble of 100 simulations.](image)
in all cases until greater than 50 nodes have failed. After that the more localized cases (σ = 0.25, σ = 0.5) break into more components. Note that the number K(t) of connected components is not strictly increasing in F(t). Occasionally the nodes of a small component (possibly a single isolated node) all fail and the component is removed from the graph completely. As the result K(t) can decrease.

Figure 3(b) shows the averaged value H(t) over all shortest paths in the graph. The initial value of H(t) for a fully connected graph should be of order Θ(1/r) ≈ 17 according to (2). Smaller H(t) indicates better network performance; as the network is damaged H(t) will increase. The Localized Scenarios with σ = 0.15 and σ = 0.25 have the worst performance with a small number of failed nodes. But for larger number of failed nodes the Localized Scenario with σ = 0.5 becomes comparable. In all cases the localized failures produce a larger (worse performance) average shortest path compared with the uniform failure scenario.

Since both connectivity and distance between all pairs of residual nodes are important, the component-averaged shortest path W(t) in Fig. 2(c) accounts for the size of a connected component as well as the average path length in the component. The value of W(t) is similar to H(t) but penalizes the growth in the number of components. This is clearly seen in the difference between the σ = 0.15 case and σ = 0.25 case. The average shortest paths H(t) are similar in both cases but with σ = 0.15 there are fewer small components compared with σ = 0.25.

Both H(t) and W(t) are based on shortest paths. Communicability Q(t), shown in Fig. 2(d), accounts for alternative paths with discounts on longer paths. In this case the larger Q(t) indicates more shorter paths and the better network. Communicability produces a different picture of the network performance. The best performing networks are with σ = 0.15 and σ = 0.25. Though the number of components and average shortest path lengths increase in both of these cases the failures are concentrated in one location which develops a hole in the center of the network with high damage and leaves an outer ring with lower damage. This damages fewer of the many redundant (not shortest) paths as in the other scenarios. All the other scenarios have very similar deterioration in communicability with respect to F(t) although the location of failures are different. It indicates that, if all possible routes between connected nodes are considered, the locations of failures have little effect on communicability in those cases.

C. Global Effects and Discrete Failures

Since networks are discrete entities and all of the measures used in this paper are discrete, averaging over a number of samples may provide results which are not achievable in any individual case. To show the kind of variance possible in individual simulations we chose the first sample of 100 simulations and show, in Fig. 3, the performance measures for some of the scenarios. Among those, scenario σ = 0.5 showed a large jump in H(t) at about 250 nodes failed. A closer look, see Fig. 4, revealed an interesting topological change. The jump happened between the damaged network with 268 and 272 failed nodes. At t = 57, with 268 failed nodes, the largest connected component forms an annulus. At t = 58 one of the four newly failed nodes breaks the ring (see the shaded area of Fig. 4). Some nodes then have to traverse many hops clockwise to reach nodes which could formerly be reached at t = 57 with only a few hops counterclockwise.

Notice that in our measures there is no monotonicity in any performance measure with respect to the magnitude of damage, i.e., value of σ. In some measures the network performance is similar in scenarios of small damage and large damage. Intuitively small damage does not cause much decrease in performance and large damage destroys networks to degenerate states in which residual network scattering and function normal as small units. For networks with different damages, e.g., between σ = 0.25 and 0.5, and between σ = 0.15 and 0.25, in our experiments multiple measures are necessary to compare the network performance.

VI. SUMMARY

In this paper we analyzed performance of a network with node failures. Uniform node failure and localized node failure were introduced as two classes of scenarios. Four network
measures were discussed to measure network performance and were used to compensate each other. Using random geometric graphs as the underlying network we simulated networks with different failure scenarios over multiple time periods and analyzed the measures. The topology of the underlying network and magnitude of failures are important factors to the network performance in the experiments. The empirical results also showed the importance of using multiple measures for assessing network performance since there is no monotonicity in any performance measure with respect to the failure scenarios.

VII. ACKNOWLEDGMENTS

This work has been supported by the Defense Threat Reduction Agency (DTRA) through grants BRCALL08-A-2-0030, and DOE Office of Science Advanced Computing Research (ASCR) program in Applied Mathematics.

REFERENCES