On the Clustering Properties of Exponential Random Networks

Nikhil Karamchandani and D. Manjunath
Dept of Elec. Engg.
IIT-Bombay, Mumbai, INDIA
nikhilk,dmanju@ee.iitb.ac.in

Srikanth K. Iyer
Dept of Math
IIT-Kanpur, Kanpur, INDIA
skiyer@iitk.ac.in

***Abstract*—**
We consider the clustering properties of one-dimensional sensor networks where the nodes are randomly deployed. Unlike most other work on randomly deployed networks, we assume that the node locations are drawn from a non uniform distribution. Specifically, we consider the exponential distribution.

We first obtain the probability that there exists a path between two labeled nodes in a randomly deployed network and obtain the limiting behavior of this probability. The probability mass function (pmf) for the number of components in the network is then obtained. We show that the number of components in the network converges in distribution. We also derive the probabilities for different locations of the components. We then obtain the probability for the existence of a k-sized component and components of size $\geq k$. Asymptotics in the number of nodes in the network are computed for these probabilities. An interesting result that we obtain is that as the number of nodes in the network tends to infinity, a giant component in which a specific fraction α of the nodes form a component almost surely does not exist for any $0 < \alpha < 1$. However, the probability converges to a non-zero value for $\alpha = 1$. Another result is that for $0 < \alpha < 1$, we can find an n_0 such that for $n > n_0$, the network almost surely does not have a giant component.

I. Introduction and Notation

Sensor nodes have limited energy and computational power. Use of distributed algorithms in executing the sensing function can optimize energy consumption. Distributed sensing algorithms are designed to consider nodes in clusters with subtasks assigned to each cluster. To enable the design of such efficient cluster-based algorithms, one needs to know the sizes of the clusters (number of nodes and geographic spread) and their relative positions in the network. In this paper, we study the clustering properties of sensor networks in which the nodes are randomly deployed. Randomly deployed networks are characterized by random geometric graphs (RGG). [8] is a classic introduction to the subject of RGGs. Study of the clustering properties of randomly deployed sensor networks is essentially the study of the connected components in the graph induced by the network, the subject of this paper.

Here, we consider one dimensional networks which can be motivated by considering a random deployment of intrusion detection sensors along a border, where we assume that the sensors are deployed from some points on the border. If the point of deployment is treated as the origin, the location of a sensor node on the border will be a random variable with a probability density function that is decreasing away from the origin. Further, the region over which the nodes get deployed would be $(0, \infty)$ (or $(-\infty, \infty)$). Thus in such deployments, we cannot assume that the sensor nodes are uniformly distributed in the sensing area, a standard assumption in most studies of wireless networks. Herein, we assume that each sensor location is drawn from an independent exponential distribution.

A. Previous Work

In this paper, we analyze finite networks and also obtain some asymptotic properties. Exact analysis of finite networks is important because the asymptotes may be approached very slowly, e.g., for a 50-node random network in $[0, 1]^2$, simulations show that the network is connected with high probability when $r = 0.31$ while asymptotic analysis suggests $r = 0.18$. Exact analysis of finite networks has been considered in [2], [3], [6]. [6] obtains limit theorems for a number of graph properties related to its connectivity, including the existence of paths and components of finite size. In this paper, we address some of these same issues for exponential networks. Some properties of exponential networks have been studied in [7]. Like in [7], we assume that the n-nodes of the one-dimensional network are distributed randomly on the positive x-axis, the node locations are i.i.d. exponential with parameter λ and all nodes have identical transmission radius r. More clustering analysis of uniform one-dimensional networks can be found in [4], [5]. Other applications for the work of this paper are in classification problems in archaeological findings, traffic light phasing, and geological problems [6].

Let $X_{(i)}$ denote the position of the i^{th} node from the origin, with $X_{(i)} < X_{(i+1)}$ for $i \in [1, n-1]$. Define $Y_i := X_{(i+1)} - X_{(i)}$ to be the distance between the i^{th} and $(i+1)^{\text{th}}$ ordered nodes from the origin, for $i \in [1, n-1]$. The pdf of Y_i, $f_{Y_i}(y_i)$ is derived follows.

$$
\begin{aligned}
\mathsf{Pr}\left(Y_i > y_i\right) &= \mathsf{Pr}\left((X_{i+1} > X_i + y_i) | (X_{i+1} > X_i)\right) \\
&= \mathsf{Pr}\left(X_{(i+1)} > y_i)\right) \\
&= \mathsf{Pr}\left(\text{min of } (n-i) \text{ exponential RVs} > y_i\right) \\
f_{Y_i}(y) &= \left((n-i)\lambda e^{-(n-i)\lambda y}\right) U(y)
\end{aligned}
$$

In deriving the above, the second equality is due to the memoryless property of the exponential distribution while the third equality is from the definition of the ordered node. The probability of there being an edge between i^{th} and $(i+1)^{\text{th}}$ ordered nodes is $\zeta_i(n) = 1 - e^{-(n-i)\lambda r}$ for all $i \in [1, n-1]$.

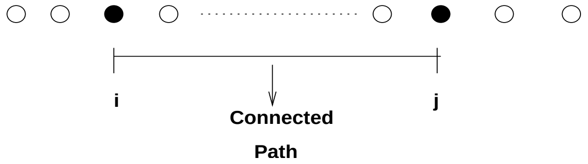

Fig. 1. Labeled nodes are at positions i and j. For there to be a path between them, each node between them should be connected to both its neighbours.

The rest of the paper is organized as follows. In Section II we obtain the probability of there being a path between two labeled nodes in a finite network and obtain its asymptotics. In Section III, we obtain the pmf for the number of components in the network. In Section IV we obtain the probability of there being a component of size k and also its asymptotics. In Section V we investigate the existence of a giant component. We conclude in Section VI.

II. Existence of a Path Between Two Labeled Nodes

In a sensor network where clusters have been assigned specific tasks, only some nodes may be made 'clusterheads'. These may be nodes with special capabilities. Thus it is important to know the existence of a path between such labeled nodes.

Let the two labeled nodes be at the i^{th} and j^{th} positions from the origin. Without loss of generality, let $j > i$. See Fig. 1 for an example. The probability of existence of a path between the two nodes is $\prod_{l=i}^{j-1} \zeta_l(n)$. Unconditioning on i and j (the labelled nodes assume position i and j with probability $\frac{(n-2)!}{n!}$) and multiplying by two to include the case $i > j$, we obtain the following property.

Property 1: The probability of there being a path between two labeled nodes is

$$P = \frac{2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\prod_{l=i}^{j-1}\zeta_l(n)}{n(n-1)}. \tag{1}$$

To obtain the asymptotics, note that P is the arithmetic mean of $\frac{n(n-1)}{2}$ terms and should be greater than the corresponding geometric mean. Thus

$$P \;>\; \left(\prod_{m=1}^{n-1}\left(\zeta_m(n)\right)^{(n-m)}\right)^{\frac{2}{n(n-1)}}. \tag{2}$$

The limit of RHS (as $n \to \infty$) of Eqn. 2 is 1. P is upper bounded by 1. Hence we have

Lemma 1: As $n \to \infty$, the limiting probability of there being a path between two labeled nodes is 1.

III. Components in the Network

A sequence of connected nodes, which are followed and preceded by a disconnected node or no nodes is called a connected component. In this section, we derive the distribution of the number of components in the network.

Let $\{\geq j\}$ denote the network of ordered nodes $j, \ldots, n$. Let $\psi_n(j,k)$, $j \in (1,\ldots,n)$, $k \in (1,\ldots,n-j+1)$ denote the probability that in an n-node network there are k components in $\{\geq j\}$. The following can be easily verified.

$$\begin{aligned}\psi_n(j, n-j+1) &= \prod_{i=j}^{n-1}(1-\zeta_i(n))\\ \psi_n(j,1) &= \prod_{i=j}^{n-1}\zeta_i(n)\end{aligned} \tag{3}$$

Note that k components in $\{\geq j\}$ can occur in one of the two ways: k components in $\{\geq (j+1)\}$ and nodes j and $j+1$ are connected or $(k-1)$ components in $\{\geq (j+1)\}$ and j not connected to $j+1$. This leads us to state the following.

Property 2: The probability that there are k components in the network, $\psi_n(1,k)$, is obtained by the following recursion.

$$\psi_n(j,k) = q_j\psi_n(j+1,k) + (1-q_j)\psi_n(j+1,k-1) \tag{4}$$

The initial conditions for the recursion will be given by Eqn. 3.

From Property 2, we observe that as $n \to \infty$, the number of components will essentially be determined by the last few nodes. To derive the limiting distribution of the number of components, consider the last node of the first component. Let $\theta_{n,m}$ denote the probability that node m is the last node of the first component in an n node network, $1 \leq m \leq n$.

For any fixed m, the probability that the last node of the first component is the m^{th} from the origin goes to 0 as $n \to \infty$, but for $m = n-s$, we can obtain the following.

$$\theta_s := \lim_{n\to\infty}\theta_{n,n-s} \;=\; \frac{Ce^{-r\lambda s}}{\prod_{i=1}^{s-1}\left(1-e^{-ir\lambda}\right)}$$

Here C is the limiting probability that the network is connected. The fact that this limiting probability exists is derived in [7]. It is also shown that $0 < C < 1$ for finite λr. As $s \to \infty$, the denominator decreases monotonically to C and θ_s goes to zero as $e^{-\lambda rs}$. To obtain k components in the network, conditional on the first component ending at $m = n-s$, we need $k-1$ components in the network composed of nodes $n-s+1, \ldots, n$. The distribution of the internodal distance between the ordered nodes $n-s+1, \ldots, n$ is exponential with parameters $(s-1)\lambda, \ldots, \lambda$. This is the same internodal distribution obtained when s nodes are distributed by choosing their distances from the origin to be exponentially distributed with parameter λ. Thus, we can write the following recursive expression for the limiting probability of the network having k components.

$$\psi(1,k) = \lim_{n\to\infty}\psi_n(1,k) = \sum_{s=k}^{\infty}\theta_s\psi_s(1,k-1)$$

$Ce^{-r\lambda s}/(\prod_{i=1}^{s-1}(1-e^{-ir\lambda}))$ decreases monotonically after a finite s, say s_1. $\psi_s(1,k)$ also decreases monotonically after a finite s, say s_2. This can be seen from Eqn. 13 that is derived later. Since the former decays to zero, the sequence for the product of individual terms of the above two sequences also decreases monotonically to zero (after a finite s_o). Hence, by ratio test, the series on the right hand side in the above equation converges to a finite value. We have thus proved the following theorem.

Theorem 1: In the exponential network, for a given λr, the number of components converges in distribution, i.e., the pmf of the number of components in the network converges as $n \to \infty$.

Next, we find the distribution of the beginning of the k^{th} component of the network. Let D_k denote the first node of the k^{th} component of the network. Clearly, $D_1 = 1$. The probability that the m^{th} node from the beginning of the network is the first node of the k^{th} component conditioned on the $(k-1)^{\text{th}}$ component beginning at m_{k-1} is given by the following.

$$\begin{aligned}
&\Pr(D_k = m \mid D_{k-1} = m_{k-1}) \\
&= \left(\prod_{i=m_{k-1}}^{m-2} \Pr(Y_i < r)\right) \Pr(Y_{m-1} > r) \\
&= \left(\prod_{i=m_{k-1}}^{m-2} \zeta_i(n)\right)(1 - q_{m-1})
\end{aligned}$$

The pmf of the beginning of the k^{th} component is now clear.

Property 3: The probability that the beginning of the k^{th} component of the network is the m^{th} node from the origin, is given by the following recursive equation.

$$\begin{aligned}
\Pr(D_k = m) &= \sum_{m_{k-1}=k-1}^{n} \Pr(D_{k-1} = m_{k-1}) \times \\
&\quad \left(\prod_{i=m_{k-1}}^{m-2} \zeta_i(n)\right)(1 - q_{m-1}) \qquad (5)
\end{aligned}$$

Next, we derive an expression for the distribution of the number of components of size m. Let $P_m^n(i,k)$ denote the probability that there are k components of size m in $\{\geq i\}$ in a network with n nodes. We are interested in $P_m^n(1,k)$. It is clear that $P_m^n(i,k) = 0$ if $mk < n - i + 1$ and

$$\begin{aligned}
P_m^n(n-m+1, 0) &= 1 - \Pr(Y_{n-m+1} < r, \ldots, Y_{n-1} < r), \\
P_m^n(n-m+1, 1) &= \Pr(Y_{n-m+1} < r, \ldots, Y_{n-1} < r).
\end{aligned}$$

Conditioning on the location of the first $j \geq i$ such that $Y_j > r$, we obtain a recursive relation for $P_m^n(i,k)$ as

$$\begin{aligned}
P_m^n(i,k) &= \sum_{j=i+1, j\neq m+i}^{n-km+1} \Pr(Y_i < r, \ldots, Y_{j-1} > r) P_m^n(j,k) \\
&\quad + \Pr(Y_i < r, \ldots, Y_{i+m-1} > r) P_m^n(m+i, k-1). \qquad (6)
\end{aligned}$$

The boundary conditions for the above recursion are

$$\begin{aligned}
&P_m^n(i,0) = \sum_{j=i}^{n-m} \Pr(Y_i < r, \ldots, Y_j > r) P_m^n(j+1, 0), \\
&P_m^n(n-km+1, k) = P_m^n(n-(k-1)m+1, k) \times \\
&\qquad \Pr\left(Y_{n-km+1} < r, \ldots, Y_{n-(k-1)m} > r\right).
\end{aligned}$$

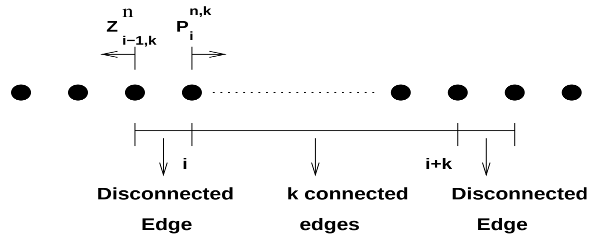

Fig. 2. A k-sized component in a n-node network.

Following the same arguments as in the proof of Theorem 1, we can show that the limiting pmf of the number of size-m components is

$$P_m(k) = \lim_{n\to\infty} P_m^n(1,k) = \sum_{s=mk}^{\infty} \theta_s P_m^s(1,k), \qquad (7)$$

where $P_m^s(1,k)$ are as given by (6).

By taking $m = 1$ in Eqn. 7, we obtain the distribution of the number of isolated nodes in the network. It is interesting to recall that for uniform networks, the probability that the network is disconnected vanishes at the same r for which the isolated nodes disappear.

IV. Existence of a k-Sized Component in the Network

The existence of a component of a particular size is an important issue because the size of clusters in the network decides what kind of load sharing happens amongst them. Define a k-sized component as a subnet of $k+1$ consecutive connected nodes, isolated from the rest of the network. Define $Z_{i,k}^n$ as the probability of the event that a k-sized component starts before or at the i^{th} node (from the origin). Let $P_i^{n,k}$ be the probability of existence of a k-sized component starting at the i^{th} node. Fig. 2 illustrates the notation. The following recursive equation is easy to verify.

$$Z_{i+1,k}^n = Z_{i,k}^n + (1 - Z_{i,k}^n) \times P_{i+1}^{n,k} \qquad (8)$$

For all $i \in [2, n-k-1]$, we can write

$$\begin{aligned}
P_i^{n,k} &= (1 - \zeta_{i-1}(n))(1 - \zeta_{i+k}(n)) \prod_{l=i}^{i+k-1} \zeta_l(n), \quad (9) \\
P_1^{n,k} &= (1 - \zeta_{k+1}(n)) \times \prod_{l=1}^{k} \zeta_l(n), \quad (10) \\
P_{n-k}^{n,k} &= (1 - \zeta_{n-k-1}(n)) \times \prod_{l=n-k}^{n-1} \zeta_l(n). \quad (11)
\end{aligned}$$

Clearly, the probability of existence of a k-sized component in the network is $Z_{n-k,k}^n$ and the boundary condition is $Z_{1,k}^n = P_1^{n,k}$. Solving the recursion, we obtain the following property.

Property 4: The probability of existence of a component of finite size k is given by

$$\xi_k(n) = Z_{n-k,k}^n = \sum_{m=1}^{n-k} \left(P_m^{n,k} \times \prod_{j=m+1}^{n-k} (1 - P_j^{n,k})\right). \quad (12)$$

To obtain the asymptotics of the above expression, we need to consider $\xi_k(n)-\xi_k(n-1)$. $\xi_k(n-1)$ can be written as follows.

$$\xi_k(n-1) \quad = \quad \sum_{m=1}^{n-k-1}\left(P_m^{n-1,k}\prod_{j=m+1}^{n-k-1}(1-P_j^{n-1,k})\right)$$

Since $\zeta_s(n-1)=\zeta_{s+1}(n)$ for all $s\in[1,n-2]$, $P_j^{n-1,k}=P_{j+1}^{n,k}$ for all $j\in[2,n-k-1]$. However, $P_1^{n-1,k}=\frac{P_2^{n,k}}{1-\zeta_1(n)}$ where, as we have defined earlier, $\zeta_1(n)$ is the probability of existence of the edge between the first and second nodes. Subtracting Eqn. 13 from Eqn. 12, we get

$$\xi_k(n)-\xi_k(n-1)=\left((1-P_2^{n,k})\times P_1^{n,k}-\frac{\zeta_1(n)P_2^{n,k}}{1-\zeta_1(n)}\right)\times\prod_{t=3}^{n-k}(1-P_t^{n,k}). \tag{13}$$

As $n\to\infty$, the difference $(\xi_k(n)-\xi_k(n-1))$ asymptotically tends to 0. Hence, we have the following result.

Lemma 2: The probability of existence of a component of size k, $\xi_k(n)$ tends to a limit as $n\to\infty$.

Now that we know $\xi_k(n)$ converges to a limit as $n\to\infty$, we show how to compute the value of this limit with a controllable accuracy. We start with the following. For $j\in[k+1,n-2]$,

$$P_{n-j}^{n,k}=e^{-(2j+1-k)\lambda r}\times\prod_{l=n-j}^{n-j+k-1}\zeta_l(n). \tag{14}$$

Note that, for all finite t, we have

$$\zeta_{n-t}(n)=1-e^{-(n-(n-t))\lambda r}=1-e^{-t\lambda r},$$

i.e., $\zeta_{n-t}(n)$ is not a function of n. Hence, for all finite j, we can denote $P_{n-j}^{n,k}$ as $P_{j,k}$. Next, as $n\to\infty$ and $j\to n$, $P_{n-j}^{n,k}\to 0$. Thus, in the expression for $\xi_k(n)$ from Eqn. 12, it is the last few terms in the summation that contribute to its value while the others tend to 0 an $n\to\infty$. Hence, if we sum terms in Eqn. 12 only for $m\in[n-j_o,n-k]$, we can write

$$\xi_k(n)==\left(\sum_{m=k}^{j_o}\left(P_{m,k}\times\prod_{j=k}^{m-1}(1-P_{j,k})\right)\right)+e^k(j_o), \tag{15}$$

where $e^k(j_o)$ is the error due to the truncation. We can bound $e^k(j_o)$ as follows.

$$\begin{aligned} e^k(j_o) \quad &< \quad \left(\sum_{i=j_o+1}^{n-1}P_{n-i}^{n,k}\right)\times\prod_{j=n-j_o}^{n-k}\left(1-P_j^{n,k}\right)\\ &< \quad \sum_{i=j_o+1}^{n-1}P_{n-i}^{n,k} \end{aligned} \tag{16}$$

From Eqn. 14, we have $P_{n-j}^{n,k}<e^{-(2j+1-k)\lambda r}$. Thus, from Eqn. 16, we note that $e^k(j_o)<\sum_{i=j_o+1}^{n-1}e^{-(2i+1-k)\lambda r}$. Then, summing up the series in RHS, we have

$$e^k(j_o)<\frac{e^{-(2j_o+1-k)\lambda r}}{e^{2\lambda r}-1}.$$

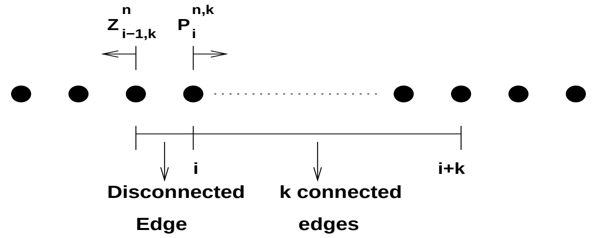

Fig. 3. A $\geq k$-sized component in a n-node network

We thus have the following lemma.

Lemma 3: The asymptotic limit for the probability of existence of a k-sized component can be computed with arbitrary accuracy as follows. For any $\epsilon>0$, we can find a corresponding finite j_o and sum only the last (j_o-k) terms in the expression for $\xi_k(n)$, as demonstrated in Eqn. 15.

A. Existence of a Component of Size $\geq k$

Redefine $Z_{i,k}^n$ as the probability that a component of size $\geq k$ starts before or at the i^{th} node (from the origin). Also redefine $P_i^{n,k}$ to be the probability of existence of a similar component starting at the i^{th} node. Fig. 3 illustrates the new notations. Clearly, the recursion form of Eqn. 8 still holds and we can write the following recursion in i.

$$Z_{i+1,k}^n=Z_{i,k}^n+(1-Z_{i,k}^n)\times P_{i+1}^{n,k}$$

For all $i\in[2,n-k]$,

$$P_i^{n,k}=(1-\zeta_{i-1}(n))\times\prod_{l=i}^{i+k-1}\zeta_l(n), \tag{17}$$

and

$$P_1^{n,k}=\prod_{l=1}^{k}\zeta_l(n). \tag{18}$$

The probability of existence of a ($\geq k$)-sized component in the network is $Z_{n-k,k}^n$. The boundary condition is $Z_{1,k}^n=P_1^{n,k}$. Along the lines of the expression in Eqn. 12, we get

$$W_k(n)=Z_{n-k,k}^n=\sum_{m=1}^{n-k}\left(P_m^{n,k}\times\prod_{j=m+1}^{n-k}(1-P_j^{n,k})\right). \tag{19}$$

We now evaluate the asymptotic limit of this probability. As in Eqn. 13, we can write

$$W_k(n)-W_k(n-1)=\left((1-P_2^{n,k})P_1^{n,k}-\frac{\zeta_1(n)P_2^{n,k}}{1-\zeta_1(n)}\right)\times\prod_{t=3}^{n-k}(1-P_t^{n,k}).$$

As $n\to\infty$, the difference goes to 0. Thus $W_k(n)$ goes to a limit as $n\to\infty$.

$P_1^{n,k}$ tends to 1 as $n\to\infty$. Denoting the asymptotic limit of all $P_i^{n,k}$ as $P_i^{\infty,k}$, the expression for $W_k(n)$ as $n\to\infty$ is

$$\begin{aligned} W_k(n)^\infty \quad &= \quad \left(\sum_{m=2}^{n-k}\left(P_m^{\infty,k}\times\prod_{j=m+1}^{n-k}(1-P_j^{\infty,k})\right)\right)\\ &+\prod_{j=2}^{n-k}(1-P_j^{\infty,k})=1, \end{aligned}$$

and hence we have the following lemma.

Lemma 4: The probability of existence of a component of size $\geq k$ goes to 1 as $n \to \infty$.

B. Existence of a k-Sized Component Starting Beyond x_o

Consider a fixed point on the positive x-axis and denote it by x_o. We obtain the probability of there being a component of size k starting at or beyond x_o. This is important because the relative positions and the distance between individual clusters in a network is crucial for maintaining communication between them. We perform the computation as follows.

Let $\mathcal{A}_i$ be the event that exactly i nodes are in $(0, x_o)$ and $\mathcal{B}_i$ be the event of the existence of a k-sized component amongst the last $(n-i)$ nodes. Clearly,

$$\mathsf{Pr}\,(\mathcal{A}_i) = (1 - e^{-\lambda x_o})^i \times (e^{-\lambda x_o})^{n-i}.$$

$\mathsf{Pr}\,(\mathcal{B}_i)$ is determined using arguments similar to those used in deriving $\xi_k(n)$ in Section IV and is

$$\mathsf{Pr}\,(\mathcal{B}_i) = \sum_{m=i+1}^{n-k} \left(P_m^{n,k} \times \prod_{j=m+1}^{n-k} (1 - P_j^{n,k}) \right).$$

This leads us to

Property 5: The probability of existence of a k-sized component starting beyond or at x_o, a given point on the positive x-axis is given by $P^{n,k}(x_o) = \sum_{i=0}^{n-k-1} \mathsf{Pr}\,(\mathcal{A}_i) \times \mathsf{Pr}\,(\mathcal{B}_i)$.

V. Existence of a Giant Component of Size $(n-1)\alpha$

For a network of n nodes, a giant component is defined as a component of size $\alpha \times (n-1)$, where $\alpha \in (0, 1]$ and is rational. As the component size has to be an integer, we have to choose a subsequence of the positive integer sequence such that each element satisfies this constraint. In the following, n corresponds to an element of this subsequence.

The probability of existence of a k-sized component in the network is $Z_{n-k,k}^n$. So, from Eqn. 12, for $k = \alpha \times (n-1)$ (ensuring that the product is an integer), we can write

$$\xi_{(n-1)\alpha}(n) = \sum_{m=1}^{n(1-\alpha)+\alpha} \left(P_m^{n,(n-1)\alpha} \prod_{j=m+1}^{n(1-\alpha)+\alpha} (1 - P_j^{n,(n-1)\alpha}) \right) \quad (20)$$

Our objective now is to derive the asymptotic value of this expression. For this, we need to compute the limit of $P_i^{n,(n-1)\alpha}$ as $n \to \infty$, for all $i \in [1, n-k]$. Beginning with Eqn. 9, substituting $k = \alpha \times (n-1)$, we get

$$P_i^{n,(n-1)\alpha} = e^{(-2n+2i+(n-1)\alpha-1)\lambda r} \times \prod_{l=i}^{i+(n-1)\alpha-1} \zeta_l(n). \quad (21)$$

Since $i \leq n - (n-1)\alpha - 1$, we can write

$$e^{(-2n+2i+(n-1)\alpha-1)\lambda r} \leq e^{(-(n-1)\alpha-3)\lambda r}.$$

The RHS of the above equation converges to 0 as $n \to \infty$. Hence, for all $i \in [2, n-(n-1)\alpha-1]$ and $\alpha \in (0,1]$, we note that $e^{(-2n+2i+(n-1)\alpha-1)\lambda r} \to 0$. Thus, $P_i^{n,(n-1)\alpha} \to 0$ as $n \to \infty$ for all $i \in [2, n-(n-1)\alpha-1]$. Similarly, Eqn. 11 can be rewritten as

$$P_{n-(n-1)\alpha}^{n,(n-1)\alpha} = e^{-((n-1)\alpha+1)\lambda r} \times \prod_{l=n-(n-1)\alpha}^{n-1} \zeta_l(n).$$

RHS in the above equation goes to 0 as $n \to \infty$ for $\alpha \in (0, 1]$ and Eqn. 10 becomes

$$P_1^{n,(n-1)\alpha} = e^{-(1-\alpha)(n-1)} \times \prod_{l=1}^{(n-1)\alpha} \zeta_l(n). \quad (22)$$

The RHS in the above equation goes to 0 for $\alpha \in (0, 1)$. Observe that $\alpha = 1$ corresponds to the case when the entire network of n nodes is connected. By substituting $\alpha = 1$ in Eqn. 22, we find that $P_1^{n,n-1} = C_n$, where C_n is given by $\prod_{i=1}^{n-1}(1 - e^{-i\lambda r})$. From [7], $C_n \to C$ as $n \to \infty$ and $0 < C < 1$. Thus expanding RHS in Eqn. 20 we can state the following result.

Theorem 2: The asymptotic limit of the probability of existence of a giant component of size $(n-1)\alpha$ is 0 for $\alpha \in (0, 1)$ and C for $\alpha = 1$.

A. Existence of a $(\geq (n-1)\alpha)$-Giant Component

We compute this probability for $\alpha \in (0, 1)$ For α = 1, the case is same as the existence of a component of size $(n-1)$ and has already been dealt above. From Eqn. 19, the probability of existence of a $\geq k$-sized component in the network is given by

$$W_k(n) = \sum_{m=1}^{n-k} \left(P_m^{n,k} \times \prod_{j=m+1}^{n-k} (1 - P_j^{n,k}) \right).$$

Substituting $k = (n-1)\alpha$ (as earlier, we choose a subsequence $\{n\}$ such that $(n-1)\alpha$ is an integer), we get

$$W_{(n-1)\alpha}(n) = \sum_{m=1}^{n-(n-1)\alpha} \left(P_m^{n,(n-1)\alpha} \prod_{j=m+1}^{n-(n-1)\alpha} (1 - P_j^{n,(n-1)\alpha}) \right).$$

From Eqn. 17 , for all $i \in [2, n-(n-1)\alpha]$

$$P_i^{n,(n-1)\alpha} = (1 - \zeta_{i-1}(n)) \times \prod_{l=i}^{i+(n-1)\alpha-1} \zeta_l(n).$$

Also, from Eqn. 18

$$P_1^{n,(n-1)\alpha} = \prod_{l=1}^{(n-1)\alpha} \zeta_l(n). \quad (23)$$

Consider $\lim_{n\to\infty} W_{(n-1)\alpha}(n)$. From Eqn. 23, $P_1^{n,(n-1)\alpha}$ goes to 1 as $n \to \infty$. Making arguments similar to that in Section IV-A, we can show that the limit evaluates to 1.

Corollary 1: As $n \to \infty$, the limiting probability of there being a $(\geq (n-1)\alpha)$-giant component is 1 for $\alpha \in (0, 1)$ and C for $\alpha = 1$.

B. Existence of a Giant Component in Large Networks

Consider a sample realization of an exponential random network, ω. We now show that if we have a network of size $n > \hat{n}(\omega)$, where $\hat{n}(\omega)$ is a finite integer, then, with probability 1, there exists no giant component in the network. To prove this, we use the Borel-Cantelli Lemma, for which we need to sum the probability of existence of a component of size $(n-1)\alpha$, over all possible n and prove that it converges to a finite limit.

Starting with Eqn. 12, $Z^n_{n-k,k} < \sum_{m=1}^{n-k} P^n_m$. Using Eqn. 9 and Eqn. 11, for all $i \in [2, n-k]$, we have $P_i^{n,k} < 1-\zeta_{i-1}(n)$, implying $P_i^{n,k} < e^{-(n-i+1)\lambda r}$. Thus,

$$Z^n_{n-k,k} < \left(\sum_{m=2}^{n-k} e^{-(n-m+1)\lambda r} + P^n_1 \right).$$

Summing up the geometric progression series in the RHS and recognizing from Eqn. 10 that $P_1^{n,k} < (1-\zeta_{k+1}(n))$, we get

$$Z^n_{n-k,k} < \frac{e^{-(k+1)\lambda r} \times (1-e^{-(n-k-1)\lambda r})}{1-e^{-p}} + e^{-(n-k-1)\lambda r}. \tag{24}$$

Let $k = (n-1)\alpha$ with α in $(0,1)$. Consider a sequence $\{n\}$ such that $(n-1)\alpha$ is an integer and sum $Z^n_{n-(n-1)\alpha,(n-1)\alpha}$ over that sequence upto ∞. Hence, using Eqn. 24 we have

$$\begin{aligned}\sum_{\{n\}} Z_{n-(n-1)\alpha,(n-1)\alpha} &< \\ \sum_{\{n\}} &\frac{e^{-((n-1)\alpha+1)\lambda r} \times (1-e^{-(n-(n-1)\alpha-1)\lambda r})}{1-e^{-p}} \\ &+ \sum_{\{n\}} e^{-(n-(n-1)\alpha-1)\lambda r}. \end{aligned} \tag{25}$$

We need to prove that the LHS of the inequality in (25) converges to a finite value. This is accomplished by showing that the RHS is finite. Taking up the second series in the RHS of (25) first, we have $\sum_{\{n\}} e^{-(n-1)(1-\alpha)\lambda r}$. Let $\hat{n}_1$ and $\hat{n}_2$ be two consecutive elements of the sequence $\{n\}$, with $\hat{n}_2 > \hat{n}_1$. The ratio of consecutive terms of this series will be $e^{-(\hat{n}_2-\hat{n}_1)(1-\alpha)\lambda r} < 1$. Hence, by ratio test, this series converges to a finite sum.

For the first series of the RHS of (25), differentiating the individual terms with respect to n, we find that derivative is > 0 for all $n < n_0$ and is ≤ 0 for all $n \geq n_0$, where $n_0 = \frac{\ln \frac{e^{p(1-\alpha)}}{\alpha}}{p(1-\alpha)}$. Hence, the terms in the series increase for elements of $\{n\}$ which are less than n_0 and decrease otherwise. As a result, the sum of the first series in the RHS of Eqn. 25 is also finite.

Therefore, RHS in Eqn. 25 converges to a finite sum and hence, so will the LHS, $\sum_{\{n\}} Z_{n-(n-1)\alpha,(n-1)\alpha}$.

Next, let the existence of a component of size $(n-1)\alpha$ (for a fixed α) in a network of n nodes, be event A_n. The probability of the occurrence of event A_n is given by $\Pr(A_n) = Z_{n-(n-1)\alpha,(n-1)\alpha}$, whose sum over all $\hat{n}$ in $\{n\}$ has been shown to be finite, implying $\sum_{\{n\}} \Pr(A_n) < \infty$. Thus, by Borel-Cantelli Lemma, with probability 1, only finitely many of the events $A_1, A_2, \ldots$ occur. This implies that if we keep increasing the number of nodes in the network, then

Theorem 3: Almost surely, for all $n > \hat{n}(\omega)$, where $\hat{n}(\omega)$ is a finite integer, there will be no giant component in the network.

VI. Discussion

We have presented a stochastic analysis of the clustering behavior of randomly deployed, one-dimensional sensor networks where the node locations are i.i.d. exponential. Clustering analysis is presented via the analysis of the components in the network. Note the two important departures from the standard literature on random sensor networks—(1) nodes are distributed in the sensing area according to a non uniform distribution and (2) in addition to asymptotic analysis, explicit formulae for finite networks have also been derived.

We first obtained the probability of existence of a path between labeled nodes. Our analysis of the number of components in the network shows that the pmf converges fairly quickly in the number of nodes. We show that the number of components converges only in distribution. We also obtained the distribution of the number of components of a given size. A probabilistic characterization of the locations of the components is also presented.

The probability that the network contains a component of specific size, its convergence and the computation of its limit with an arbitrary accuracy have also been discussed. We showed that for large networks, a component with a specified minimum size exists with probability one. Also, while we cannot find a component in which a specific fraction of nodes is connected, with probability one, we can almost surely find a component in which at least a specified fraction form a component.

References

[1] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *A First Course in Order Statistics,* John Wiley and Sons, New York (1993).

[2] M. P. Desai and D. Manjunath. On the connectivity of finite ad hoc networks. IEEE Communications Letters. **6**(10):437–439 (2002).

[3] M. P. Desai and D. Manjunath, "On distances and random geometric graphs," Submitted (2004). http://www.ee.iitb.ac.in/ dmanju/Publns/mpd2.ps.gz.

[4] E. Godehardt, Graphs as Structural Models: The Applications of Graphs and Multigraphs in Cluster Analysis, Vieweg, Braunschweig and Wiesbaden, (1990).

[5] E. Godehardt, Probability models for Random Multigraphs with Applications in Cluster Analysis. Annals of Discrete Mathematics. **55**:93-108 (1993).

[6] E. Godehardt and J. Jaworski, On the Connectivity of a Random Interval Graph. Random Structures and Algorithms. **9**:137–161 (1996).

[7] Srikanth K. Iyer, D. Manjunath and Bhupendra Gupta, "On the Topological Properties of a Class of Non Uniform Sensor Networks in One Dimension," *Preprint,* 2004.

[8] M. D. Penrose. Random Geometric Graphs, Oxford University Press (2003).