Connectivity for Multiple Multicast Trees Schemes in Ad Hoc Networks

Wei Wei and Avideh Zakhor
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley, CA 94720, USA
Email: {weiwei, avz}@eecs.berkeley.edu

***Abstract*— In this paper, we first define level of tree connectivity for multicast routing schemes in wireless ad hoc networks. Then given the level of tree connectivity, we examine the relation between node density required for single and double tree schemes both analytically and experimentally. For high density wireless ad hoc networks and a given level of tree connectivity, our analysis shows that the difference between required node densities for the two schemes is no larger than a logarithm factor of double tree scheme.**

I. INTRODUCTION

Wireless ad hoc networks are temporary wireless networks without an infrastructure, formed by multiple wireless mobile nodes via peer-to-peer communication. This implies that it is not always possible to keep all the nodes connected.

Multicast is an essential technology for many applications, such as group video conferencing and video distribution, and it results in bandwidth savings compared to multiple unicast sessions. Due to mobility of wireless nodes, topology of ad-hoc networks changes frequently. Thus the established multicast tree is likely to be broken during the period of the transmission, causing interruption, pause, or jerkiness in the quality of received video signal. As such, tree diversity is an effective technique to reduce the effects caused by tree failure; however, it also reduces connectivity level as compared to single tree scheme. Specifically, if double tree scheme requires a significant increase in node density in order to keep a high connectivity level, it may be too expensive to implement in practical situations.

In this paper, we explore tree connectivity for multiple disjoint multicast trees schemes in wireless ad hoc networks. We define the level of tree connectivity as the expected value of the ratio between the average number of receivers connected to each multicast tree and the total number of receivers. Specifically, for a given level of tree connectivity, we study the relation between node density required for single and double tree schemes. For high density wireless ad hoc networks, our analysis shows that the difference between node densities of two schemes is no larger than a logarithm factor of node density of double tree scheme; this means that the required density for double tree scheme is not significantly larger than that of single tree scheme, and that tree diversity is a feasible technique to improve the robustness of multicast video transmission over wireless ad hoc networks.

II. RELATED WORK AND PROBLEM FORMULATION

In recent years, there has been a number of papers on the problem of path connectivity in wireless ad hoc networks [1-5]. The problem can be described as follows: If two nodes were to be chosen at random, what is the probability of having a path between them? Let r denote the radio link range, λ the density of nodes, i.e. in the number of nodes per unit area, n the total number of nodes, and m the total number of receivers. We assume a Poisson Boolean model $B(\lambda, r)$, where nodes are distributed according to a Poisson point process of density λ in a square field of size A.

Philips et. al. [1] have shown that choosing $r = \sqrt{\frac{(1-\epsilon)\ln(A)}{\pi\lambda}}$ for any $\epsilon > 0$ leads to an asymptotically disconnected network as $A \to \infty$. They conjectured that choosing $r = \sqrt{\frac{(1+\epsilon)\ln(A)}{\pi\lambda}}$ would lead to asymptotic connectivity. Gupta and Kumar [2] have proved that $r = \sqrt{\frac{\ln(n)+c(n)}{\pi\lambda}}$ leads to asymptotic connectivity if and only if $c(n) \to \infty$ as $n \to \infty$. Dousse et. al. have stated that there exists one critical density λ_c, such that if the density $\lambda < \lambda_c$, all the connected clusters are almost surely bounded; otherwise, almost surely there exists one unique unbounded super connected cluster [3].

In this paper, we study tree connectivity of a scheme, which uses two disjoint multicast trees. The problem is formulated as follows: given tree connectivity level P, what is the relation between D_1 and D_2, the required minimal node density for single and double disjoint multicast tree schemes respectively, in order to ensure tree connectivity level of P?

We first define tree connectivity P as follows:

$$P \triangleq \frac{E[N]}{M} \tag{1}$$

where M is the product of the total number of receivers and the number of disjoint trees. $N = \sum_{i=1}^{m} n_i$, with n_i denoting the number of disjoint trees that receiver i connects to, which can be 0, 1 or 2. Given a random topology with n nodes, one random sender and m random receivers, N denotes the sum of receivers connecting to each multicast tree, and $E[N]$ is the expected value of N over all topologies. Intuitively, if the node density is low and ratio range is small, it is difficult to construct trees that connect to all m nodes, and hence tree connectivity is expected to be low. On the other hand, for sufficiently high node density and sufficiently large radio range, we would expect the connectivity to be close to one, i.e. to be able to construct trees that connect to almost all the receivers.

We make following assumptions:

1) We assume that node density of the network is higher than the critical density λ_c, i.e. the network is in the supercritical phase, and there exists one infinite unbounded super-cluster H in the network. We use θ to denote the fraction of nodes belonging to the unbounded super-cluster. θ can also be defined as the probability of an arbitrary node belonging to the unbounded cluster, and is called percolation probability. Note that in the supercritical phase, θ is close to 1.
2) The total number of nodes is infinite, and the size of the deployment area is infinite.
3) The ratio of the number of receivers to the total number of nodes is very small.

III. Analysis

A. Single Multicast Tree Construction Schemes

We have the following result for single tree schemes.

Theorem 1: The tree connectivity level for a single tree scheme is given by

$$P_1 \approx \theta_1^2 \tag{2}$$

where θ_1 denotes the fraction of nodes belonging to the super-cluster in the single tree case.

The proof is shown in the Appendix I.

B. Two Disjoint Trees Construction Schemes

We establish the following result for two disjoint trees schemes.

Theorem 2: Assume a given tree connectivity level P. There exists at least one double disjoint tree scheme whose required node density D_2 to achieve P satisfies

$$D_2 - \frac{\ln(\pi D_2 r^2 + 1)}{\pi r^2} \leq D_1 \leq D_2 \tag{3}$$

where D_1 is the required node density to achieve P in a single tree case.

Proof: We use S to denote the sender, R_i to denote Receiver i, $i = 1, \ldots, m$, m and n to denote the number of receivers and nodes respectively, and H and H' to denote the connected infinite supercluster and the one after removing all the middle nodes of the first tree respectively. From Equation (1), tree connectivity of a double tree scheme is expressed as $P_2 = E[\frac{N}{2m}]$, where $N = \sum_{i=1}^{m} n_i$, with n_i denoting the number of disjoint trees that receiver i connects to. We define N_1 and N_2 as the number of receivers which belong to H and H' respectively. The fraction of nodes belonging to H is $\theta_2 \approx 1 - \exp(-\pi D_2 r^2)$ [6].

Similar to single tree scheme proof in the Appendix I, $E[N|S \notin H]$ is negligible compared to $E[N|S \in H]$, so we only consider $E[N|S \in H]$. By definition of N_1 and N_2,

$$N = N_1 \times 1(S \in H) + N_2 \times 1(S \in H') \tag{4}$$

Similar to Equation (19) in Appendix I,

$$E[N_1 \times 1(S \in H)] = m\theta_2^2 \tag{5}$$

We also have

$$\begin{aligned} &E[N_2 \times 1(S \in H')] \\ &\quad = E[1(S \in H')]E[\sum_{i=1}^{m} 1(R_i \in H')] \end{aligned} \tag{6}$$

Since $S, R_1, \ldots R_m$ are independent identically distributed, $P(S \in H') = P(R_i \in H'), i = 1, \ldots m$. We denote this probability as $P(R \in H')$, and simplify Equation (6) as

$$\begin{aligned} E[N_2 \times 1(S \in H')] &= mP(S \in H')P(R \in H') \\ &= mP^2(R \in H') \end{aligned} \tag{7}$$

We have the following claim.

Claim 1: define receiver sets $A = \{R : \text{Ne(R)} \geq 2, R \in H\}$ and $B = \{R : R \in H'\}$, where Ne(R) denotes the number of receiver R's neighbors, H and H' denote the connected infinite supercluster and the one after removing all the middle nodes of the first tree respectively. Then there exists at least one tree scheme that makes $A \approx B$.

The proof is shown in the Appendix II.

From the Claim, we get:

$$\begin{aligned} &P(R \in H') \\ &\quad \approx P(\text{Ne(R)} \geq 2, R \in H) \\ &\quad = P(\text{Ne(R)} \geq 2) - P(\text{Ne(R)} \geq 2, R \notin H) \end{aligned} \tag{8}$$

From the definition of point poisson process, we have

$$P(\text{Ne(R)} = 0) = \exp(-\pi D_2 r^2) \tag{9}$$

$$P(\text{Ne(R)} = 1) = \pi D_2 r^2 \exp(-\pi D_2 r^2) \tag{10}$$

Thus,

$$\begin{aligned} &P(\text{Ne(R)} \geq 2) = 1 - P(\text{Ne(R)} = 0) - P(\text{Ne(R)} = 1) \\ &\quad = 1 - \exp(-\pi D_2 r^2) - \pi D_2 r^2 \exp(-\pi D_2 r^2) \end{aligned} \tag{11}$$

Also,

$$\begin{aligned} &P(\text{Ne(R)} \geq 2, R \notin H) \\ &\quad = P(\text{Ne(R)} \geq 2 | R \notin H) \cdot P(R \notin H) \\ &\quad = P(\text{Ne(R)} \geq 2 | R \notin H) \cdot \exp(-\pi D_2 r^2) \\ &\quad \ll \exp(-\pi D_2 r^2) \end{aligned} \tag{12}$$

The last inequality holds because in a high density network, when a node does not belong to the super-cluster, very likely it is isolated, i.e. has no neighbors [6]. Substituting (11) and (12) into (8), we get:

$$\begin{aligned} &P(R \in H') \\ &\quad \approx 1 - \pi D_2 r^2 \exp(-\pi D_2 r^2) - \exp(-\pi D_2 r^2) \end{aligned} \tag{13}$$

And from (4), (5), and (7), we get:

$$\begin{aligned} P_2 &= E[\frac{N}{2m}] \\ &= \frac{\theta_2^2}{2} + \frac{P^2(R \in H')}{2} \\ &\geq P^2(R \in H') \end{aligned} \tag{14}$$

Letting $P_1 = P_2$, and using Equation (2), we get $P_1 = \theta_1^2 = P_2 \geq P^2(R \in H')$. Substituting Equation (13) for $P(R \in H')$, we get:

$$(\pi D_2 r^2 + 1)\exp(-\pi D_2 r^2) \geq \exp(-\pi D_1 r^2) \quad (15)$$

Therefore,

$$D_2 - \frac{\ln(\pi D_2 r^2 + 1)}{\pi r^2} \leq D_1 \leq D_2 \quad (16)$$

■

Thus the difference between D_1 and D_2 is only a logarithm factor of D_2. Since both $D_1 \to \infty$ and $D_2 \to \infty$ as $n \to \infty$ in order to keep connectivity of the network [1][2], the difference, which is a logarithm factor of D_2, is small compared to the value of D_1 and D_2.

IV. Numerical Results

Our theoretical analysis is based on a series of ideal assumptions. In this section, we conduct numerical simulations to verify our analysis by investigating tree connectivity in a quasi-ideal region. We distribute nodes according to a two dimensional Poisson point process over a square area. The radio range is 25 pixels, and the size of the area depends on the number of total nodes and node density. The following simulations are repeated 5000 times to obtain average values.

The representative for single multicast tree schemes is a shortest path tree scheme based on routing message flood [7]. This scheme is optimal since it is able to build a tree to connect all the receivers that actually connect to the sender. We design the representative for double disjoint multicast tree in a serial fashion as follows. We first build a shortest path multicast tree. Then after requiring all the middle nodes in the first tree not to be middle nodes of the second tree, we build another disjoint shortest path tree. We expect the tree connectivity to be approximately optimal if the number of middle nodes of the first tree is minimal. Since the number of middle nodes of a shortest path tree is approximately minimal, the result achieved by our proposed multiple trees multicast routing scheme approximates the optimal one.

Our first simulation investigates difference in tree connectivity between single and double tree schemes as a function of node density. In this simulation, The number of receivers is set to be $0.1 \times \sqrt{n}$, where n is the number of total nodes. Figure (1) compares tree connectivity in the single and double tree schemes for a 1000 node network. As seen, tree connectivity of both schemes increases with node density, but the curve corresponding to double tree is below that of single tree. Specially, there is a performance gap between the two schemes when node density is between 2×10^{-3} and 6×10^{-3} nodes per unit area. Since our specific double tree scheme is sub-optimal, we expect the performance of optimal double tree scheme to be superior to that shown in Figure (1), resulting in a smaller performance gap. From *Theorem 2*, we deduce that the difference in tree connectivity between single and double tree schemes to be very small in a high density network. This is confirmed in the simulations shown in Figure (1).

Figure (2) shows simulation results relating the required node densities D_1 and D_2, given the same level of tree connectivity. It also shows corresponding lower bounds and upper bounds for node density D_1 given by *Theorem 2*. From Figure (2), we can see that simulation results fit in between the lower and upper bounds quite well in high density region, i.e. beyond 5×10^{-3} nodes per unit area. More interestingly, D_1 also fits into the region predicted by Equation (16) in low density region.

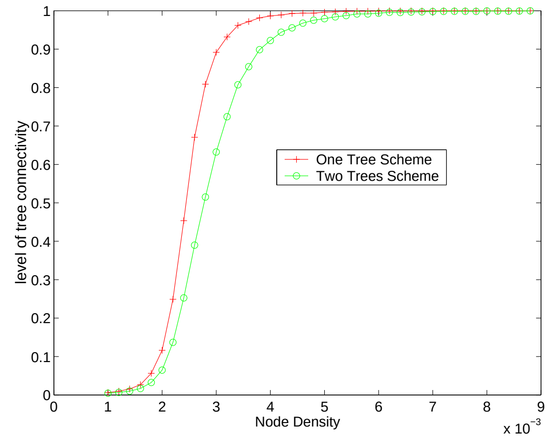

Fig. 1. Comparison of tree connectivity of one tree scheme and two disjoint trees scheme with receiver ratio 0.1

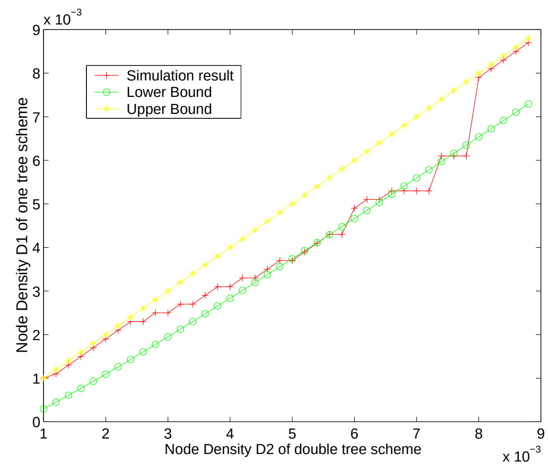

Fig. 2. Comparison of node densities of two kinds of schemes with receiver ratio 0.1

V. Parallel Two Trees Scheme Versus Serial Two Trees Scheme

In the previous section, we proposed a serial double tree scheme, which builds two disjoint trees serially. This scheme achieves good tree connectivity, but at the same time has some undesirable properties in practice. Routing overhead and delay of the serial scheme are twice as much as that of a parallel scheme that builds two trees simultaneously. This is not suitable for many delay sensitive applications, such as video.

In this section, we propose a parallel double tree scheme, which builds two nearly disjoint trees simultaneously in a distributed way, and also achieves reasonably good tree connectivity. In the parallel scheme, all the nodes are classified into one of three categories: general nodes, group 0 nodes, and group 1 nodes. If the number of a node's neighbors is smaller than a threshold, it classifies itself as a general node. Otherwise, it randomly classifies itself as either group 0 or group 1 with probability of $\frac{1}{2}$. The sender broadcasts to the entire network a member JOIN message. When a node receives a non-duplicate JOIN message, if it is not a general node, it stores the upstream node ID and rebroadcasts the message if the upstream node is either in the same group, or is a general node, or the sender; otherwise if the node is a general node, it simply stores the upstream node ID and rebroadcasts the non-duplicate JOIN message. The process continues, until in the end, the scheme builds two trees simultaneously. The first tree consists of group 0's nodes and some general nodes, and the second one consists of group 1's nodes and some general nodes. It is possible that these two trees share some general nodes, but the ratio of general nodes is very small in a high density network. Thus the scheme builds two nearly disjoint trees in one round. Although the scheme can build two entirely disjoint trees by eliminating the concept of general nodes, the classification of nodes into two groups is equivalent to dividing the network into two networks with half the node density, thus resulting in very poor tree connectivity. General nodes act as bridges between high density areas, enhancing tree connectivity a lot at the expense of loss of disjointness.

The above process results in the distribution of group 0 nodes and group 1 nodes not to be locally uniform, i.e. the number of group 0 nodes and group 1 nodes are not approximately equal in a small local area. This could potentially reduce tree connectivity of the parallel scheme. To alleviate this, we apply an equalization scheme to make its distribution as uniform as possible. Each node is able to obtain its neighbors' ID and type from periodic beacon messages. Without loss of generality, we assume the number of group 0's nodes is larger than that of group 1's nodes. Each node counts the number of nodes which belong to group 0 and have a smaller ID than itself. This number is denoted as pos. A node changes to group 1, if it belongs to group 0, and the following condition is satisfied.

$$(num\_of\_group0\_nodes - pos) > num\_of\_group1\_nodes \tag{17}$$

By applying this scheme, the distribution of two groups' nodes is locally balanced in a distributed manner.

We compare tree connectivity of the parallel scheme and the serial scheme using simulations. The number of receivers is $0.1 \times \sqrt{n}$, where $n = 1000$ is the total number of nodes. We select the threshold to be 8 neighbors. As shown in Figure (3), for node densities between 2×10^{-3} and 4.8×10^{-3}, tree connectivity of the parallel scheme is higher than that of the serial scheme; this is primarily because there are a large number of overlapped nodes between two trees obtained by the parallel scheme. When node density is above 4.8×10^{-3} nodes per unit area, tree connectivity of the serial scheme becomes slightly higher than that of the parallel scheme. Simulation results also show that, tree disjointness of the parallel scheme is only slightly worse than that of the serial scheme, when the node density is higher than 5×10^{-3} nodes per unit area.

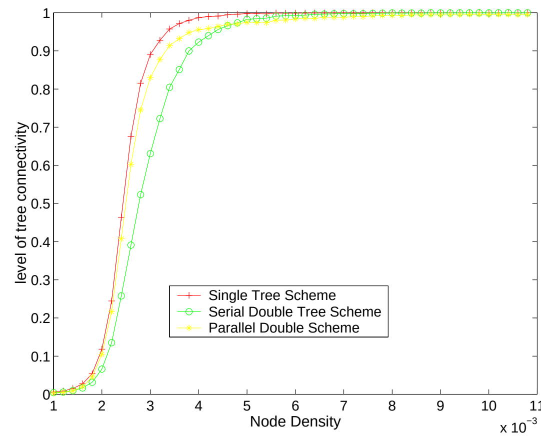

Fig. 3. Comparison of tree connectivity of the parallel scheme and the serial scheme

APPENDIX I
PROOF OF *Theorem 1*:

S, $R_i, i = 1, \ldots, m$, and H are defined as in section III.B. We use $X(S)$ to denote the number of nodes of the cluster to which the sender belongs. By the definition of tree connectivity level, we get:

$$P_1 = E[\frac{N}{m}]$$

where N denotes the number of receivers connecting to the tree, and $M = m$ for single tree schemes. By the property of conditional expectation, we further get

$$\begin{aligned} P_1 &= \frac{1}{m}\{\theta_1 \times E(N|S \in H) \\ &\quad +(1-\theta_1) \times E(N|S \notin H)\} \end{aligned} \tag{18}$$

where θ_1 denotes the fraction of nodes belonging to the supercluster in the single tree case. Under the assumption of a dense ad hoc network, it has been shown that $\theta_1 \approx 1 - \exp(-\pi D_1 r^2)$ [6], where D_1 is node density of the network.

Using the definition of $N = \sum_{i=1}^{m} n_i$, we get:

$$\begin{aligned} E(N|S \in H) &= E(\sum_{i=1}^{m} n_i | S \in H) \\ &= m\theta_1 \end{aligned} \tag{19}$$

We now show that $E[N|S \notin H] \to 0$ as $n \to \infty$. Using the definition of conditional expectation, we get:

$$E(N|S \notin H) = \sum_{i=1}^{m} iP(N = i|S \notin H) \tag{20}$$

Besides the infinite size supercluster H, there are infinite number of finite size small clusters in the network. Let k denote the number of nodes of the largest finite cluster, and $X(S)$ denote the number of nodes of the cluster $C(S)$ to which the sender S belongs. Note that when $X(S) \leq i$, at most there are $i-1$ receivers belonging to $C(S)$, thus

$P(N = i|S \notin H, X(S) \leq i) = 0$. Then Equation (20) can be further expressed as

$$E(N|S \notin H) = \sum_{i=1}^{m}\{i \sum_{j=i+1}^{k} [P(N = i|S \notin H, X(S) = j) \cdot P(X(S) = j)]\} \quad (21)$$

Also from [6], and the fact that $k \ll D_1$, we have

$$\begin{aligned} P(X(S) = j) &= \exp\left[-4\pi r^2 D_1 + (j-1)\ln\left(\frac{D_1}{j-1}\right)\right] \\ &< \exp\left[-4\pi r^2 D_1 + (k-1)\ln\left(\frac{D_1}{k-1}\right)\right] \\ &= P(X(S) = k) \end{aligned} \quad (22)$$

Taking into account Equation (22), Equation (21) can be simplified as

$$E(N|S \notin H) < \sum_{i=1}^{m}[i \sum_{j=i+1}^{k} P(N = i|S \notin H, X(S) = j)] \cdot P(X(S) = k) \quad (23)$$

It is easy to show that

$$P(N = 1|S \notin H, X(S) = j) \approx \frac{m(j-1)}{n} \quad (24)$$

and

$$2P(N = 2|S \notin H, X(S) = j) \approx \frac{m(m-1)(j-1)^2}{n^2} \quad (25)$$

given that $j \leq k \ll n$. By assumption 3, $m \ll n$, and comparing the right hand side of Equations (24) and (25), we get:

$$2P(N = 2|S \notin H, X(S) = j) \ll P(N = 1|S \notin H, X(S) = j) \quad (26)$$

From (26) and by induction, we get:

$$(i+1) \sum_{j=i+2}^{k} P(N = i+1|S \notin H, X(S) = j) \ll i \sum_{j=i+1}^{k} P(N = i|S \notin H, X(S) = j) \quad (27)$$

where $i = 1, 2, \ldots, m-1$.

Considering Equation (27) and (23), we get:

$$E(N|S \notin H) < m \sum_{j=2}^{k} P(N = 1|S \notin H, X(S) = j) \cdot P(X(S) = k)$$

Substituting Equation (24), we get:

$$E(N|S \notin H) < m \sum_{j=2}^{k} \frac{m(j-1)}{n} \cdot P(X(S) = k) \quad (28)$$

By assumption, $m \ll n$ and $k \ll n$, and $D_1 \to \infty$ in order to keep connectivity as $n \to \infty$ [2]. Therefore, using Equation (22), we conclude that $P(X(S) = k) \to 0$ as $D_1 \to \infty$. From Equation (28), we obtain

$$E(N|S \notin H) \to 0 \quad (29)$$

as $n \to \infty$.

By the dense network assumption, $\theta_1 \approx 1$. Taking into account Equations (18), (19), and (29), we get:

$$P_1 \approx \theta_1^2 \quad (30)$$

APPENDIX II
PROOF OF CLAIM 1:

Proof:

It is easy to show that $B \subseteq A$. We only need to show that $A \subseteq B$ approximately. Without loss of generality, we assume that the sender belongs to set H' and is well connected. A well connected node has so many neighbors that losing a few of them does not affect its connectivity status. If not, we can select a well connected node belonging to set H', and build a core based tree[8].

We construct the first tree in such a way that each receiver belonging to set A is only a leaf node, which means that each receiver only connects to one neighbor in the first tree. We define $K = \{$ middle nodes of the first tree $\}$ Now we want to show there is only approximately one neighbor of each receiver belonging to set K. We pick an arbitrary receiver R belonging to set A. By assumption 3, the distribution of receivers is scattered, so quite likely there are no other receivers, which are in the two-hop neighborhood of receiver R. Thus quite likely receiver R does not share any neighbors with other receivers, and hence there is approximately only one neighbor of receiver R belonging to K. Thus after removing all the middle nodes of the first tree, receiver R still has at least one neighbor left. It has been shown in [3] that for high density networks, the number of independent paths from receiver R to the sender is equal to $Ne(R)$. Thus receiver R still has at least one path connecting to the sender, which belongs to set H'. Therefore $R \in H'$ approximately. Thus $A \approx B$. ∎

ACKNOWLEDGMENTS

The authors would like to thank Olivier Dousse (EPFL), for the helpful discussion on the analysis of tree connectivity of two disjoint trees schemes.

REFERENCES

[1] T. Philips, S. Panwar, A. Tantawi. "Connectivity Properties of a Packet Radio Network Model", *IEEE Trans. On Information Theory*, Sep. 1989

[2] P. Gupta, and P. Kumar. "Critical Power for Asymptotic Connectivity in Wireless Networks". *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming* 1998

[3] O. Dousse, P. Thiran and M. Hasler, "Connectivity in ad-hoc and hybrid networks", *IEEE Infocom 2002*

[4] D. Yu, and H. Li. "On the Definition of Ad Hoc Network Connectivity", *ICCT 2003*

[5] L. Shu, M. Desai, and R. Mangoubi. "Average Connectivity Properties of Wireless Ad Hoc Networks." *WCNC 2003*

[6] R. Meester and R.Roy. "Continuum percolation", *Cambridge University Press, Cambridge*, 1996

[7] S. Lee, W. Su, and M. Gerla. "On-Demand Multicast Routing Protocol in Multihop Wireless Mobile Networks", *ACM/Kluwer Mobile Networks and Applications*, 2000

[8] T. Ballardie, P. Francis, and J. Crowcroft. "Core based trees (CBT) an architecture for scalable inter-domain multicast routing", *Computer Communication Review*, Oct. 1993