Self-connectivity Estimation for Super Node Overlay Creation in Ad-hoc Networks

Jagadish Ghimire
Institut TELECOM SudParis

Mehdi Mani
Institut TELECOM SudParis

Teerapat Sanguankotchakorn
Asian Institute of Technology

Noël Crespi
Institut TELECOM SudParis

***Abstract*—In this paper we propose a connectivity index that allows each node to estimate how well it is connected to the rest of the network. We use this index to select the supernodes (SNs) that create a P2P overlay over ad hoc network. We show that applying this index to select the SNs, despite its simplicity, leads to a high robustness in the overlay network and decrease the overlay disconnection probability. We evaluate our connectivity index via extensive simulation in ns-2 .**

I. Introduction

Deployment of P2P systems on the ad-hoc networks brings in interesting services for local use cases and disaster rescue scenario where there is no network infrastructure and internet access. Among the different architectures, P2P supernode (SN) overlay fits better to the ad-hoc network conditions. Skype [1]is an example of supernode overlay system. In a supernode based overlay networks, only a fraction of capable nodes assume the role of supernode and only supernodes together form an overlay network. All other nodes associate with one (or more) supernode(s) and rely on the supernodes for all distributed operation including destination lookup and their own advertisement. In [2], it is mentioned that such hierarchy is very important in wireless ad hoc networks to account for the heterogeneity of nodes, scalability of the algorithm and the impact of node mobility.

Connectivity of supernodes to the rest of the network plays an important role in the stability of the overlay in an ad-hoc network. Therefore, to establish an overlay, it is important to select supernodes from a set of nodes with robust connectivity. Unlike in wired network, wireless networks exhibit poor stability of the topology. Due to this, link failure is much common in such network. Any node with very few connections to the rest of the network is thus left vulnerable. In case of link failure, such nodes are likely to get offline from the whole network. This problem is amplified if the node which was segregated from the network was a supernode. This will be a disaster in the sense that all the client nodes registered to this particular supernode will be offline despite the fact that they had an existent underlay connectivity to the rest of the nodes.

Quantifying connectivity of a node in a given network is challenging. There have been work including [3], [4] and [5] which are concerned about network connectivity. [6] studies different existing connectivity metrices for studying connectivity properties of a network for detecting connectivity attacks. Very similar objective is considered in [7] which studies the Randic Connectivity Index (from which our proposed idea is derived) of an E-R Random Graph and tries to show how changes in network connectivity can be detected. Both of these work deal with the quantification of network connectivity. Note that our objective is quite different in the sense that we want to use the proposed index in quantifying node connectivity. To the best of our knowledge, there is not considerable work in quantification of node connectivity.

In this paper, the objective is to define a connectivity index for a node to estimate its connectivity stability without having the whole network information. We move on to use the proposed connectivity index as a criteria of supernode selection and show how such connectivity information is beneficial in the context of supernode selection in P2P systems over wireless ad hoc network.

II. The Proposed Connectivity Index

Our proposed connectivity index derives its idea from the field of molecular chemistry. Randic (in 1975) proposed a so-called “Randic Connectivity Index (CI)” as an index to express the branching (or connectedness) of a molecular structure [8]. We intend to use the index of molecular branching into our network (viewed as an undirected weighted graph). Important, however, is to note that the original index is defined to represent the connectedness of a whole graph. In contrast, our objective is to propose an index that could quantify connectivity in node level. Here, we intend to propose a simple “one-valued” connectivity index of a node that could quantify the connectivity of a node in a given network.

A. Randic Connectivity Index

For a given unweighted graph G, Randic Connectivity Index (CI) is defined in [8] as:

$$\chi(G) = \sum_{u,v \in V(G)} \frac{1}{\sqrt{d(u).d(v)}} \tag{1}$$

where $d(u)$ and $d(v)$ represent the degree of vertex u and v respectively, and $V(G)$ represent the vertex set of the graph.

Randic Connectivity Index definition is extended for general weighted undirected graph as follows:

$$\chi(G) = \sum_{u,v \in V(G)} \frac{w_{u,v}}{\sqrt{d(u).d(v)}} \tag{2}$$

where $w_{u,v}$ is the weight of the edge adjacent to vertices u and v. Also, we note here that $0 \leq w_{u,v} \leq 1$.

B. From Network-connectivity to Node-connectivity

Randic Index is the index of connectivity of the whole network (or graph). However, we are interested in quantifying the connectivity of a node in a given network. The common approach taken is to quantify connectivity by limiting the view of network to a local level (for example, number of one-hop neighbors used as the measure of node-connectivity). Our approach is to somehow associate a sub-graph corresponding to each node, and find the Randic based CI of the sub-graph. The calculated CI of the subgraph associated with a node can be regarded as the connectivity value of the node itself. By doing so, our proposed index intends to capture the inherent trade-off that exists between cost and accuracy. We demonstrate our concept with an example.

For example, we show the actual graph and the associated sub-graphs of some vertices of the graph in Figures 1 and 2. The criteria, used in this example is a 2-hop criteria wherein a vertex is associated with a sub-graph such that it includes all the nodes that are maximally 2-hops away from the vertex.

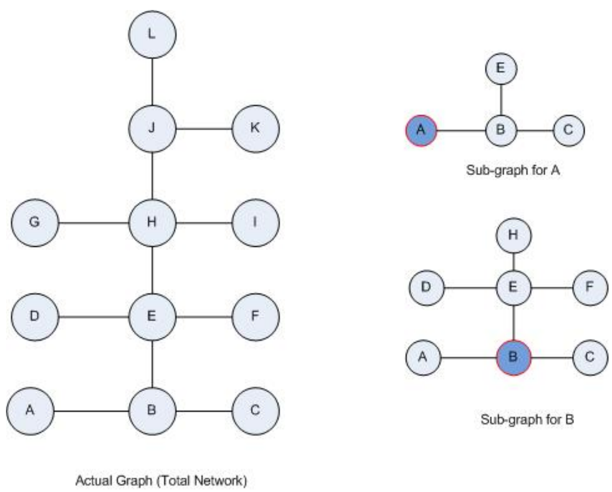

Fig. 1. Illustration of calculation of CI

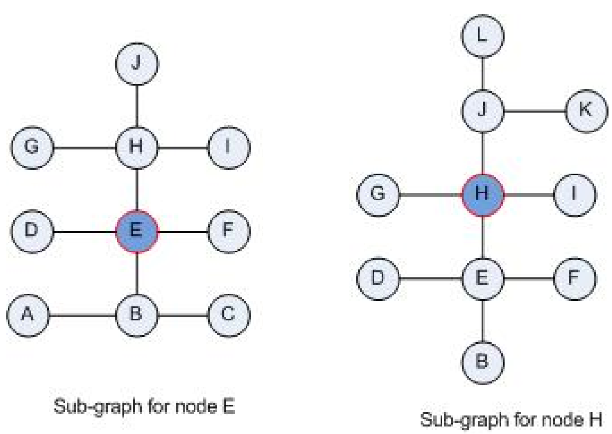

Fig. 2. Illustration of calculation of CI contd..

In Figure 1, we can see the actual network graph represented as a labeled graph with 12 vertices and 12 edges. In Figure 1, we also see the sub-graphs corresponding to node A and node B. We considered 2-hops for constructing a sub-graph. So, we associate a simple rule such that each node has a sub-graph rooted at it. Sub-graph for node E and H (with 2-hop criteria) are shown in Figure 2.

Now after we associate a sub-graph to each node, we calculate the CI of the sub-graph (as already illustrated). For example, Connectivity-value of

Node A: 1.73205
Node B: 2.94338
Node E: 4.19338
Node H: 4.19338

We see roughly how this CI is correlated to the level of branching of the graph. To illustrate this, we imagine two nodes with their own associated sub-graphs such that they have the same number of 2-hop neighbors but with a different level of branching as shown in the Figure 3. If we calculate the CI of the two sub-graphs, we find that CI of node A is 5.06538 where as CI of node B is 5.7735. Clearly, it shows that CI favors more branched nodes which is desirable for our purpose.

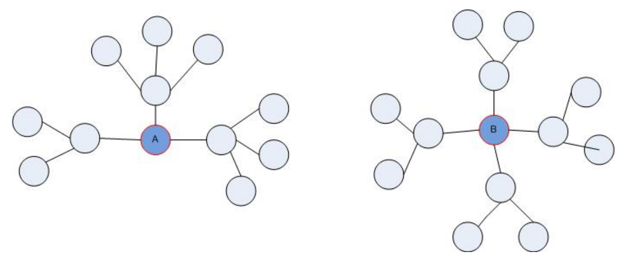

Fig. 3. Illustration of calculation of CI contd..

C. CI to account link quality

Since our index should be able to reflect the connectivity due to link-quality also, we have to consider the link-qualities of links in CI calculation. Link quality is considered as the weight of the corresponding arc in a weighted graph. Packet delivery ratio (PDR) can be considered an example for such weights of the links.

If we limit our definition of connectivity to 2-hops, CI of any node v is calculated as shown in equation 3.

$$CI(v) = \sum_{j \in 2-hop \text{ neighborhood of } v} \frac{w_{v,j}}{\sqrt{d(v).d(j)}} \tag{3}$$

Where $w_{v,j}$ represents the weight of the link (v, j). Weights are the reflections of link-quality such that it is between 0 and 1; 0 corresponding to non-existent link where as 1 represents a link with maximum link quality.

III. Simulation Model and Performance Metric

In this section we present the simulation setup for the evaluation of the proposed connectivity index (CI). The simulation have been carried out in ns-2 [9]. Our aim is to answer how the node connectivity based supernode election affects the robustness of the overlay network.

To evaluate the robustness of the overlay we define the Overlay Disconnection Rate (ODR) as our performance metric. Then we apply our proposed CI for selection of supernodes and evaluate the robustness such overlay using ODR. Furthermore, we consider three more strategies for selection supernodes and compare their influence with our proposed CI. In the following we present more details about ODR and the considered supernode selection strategies.

A. Overlay Disconnection Rate (ODR)

Overlay Disconnection Rate (ODR) is defined as the rate at which a given overlay is disconnected, subjected to the link/node failure. An overlay is regarded as disconnected if one or more supernodes (along with the associated CNs) lose an overlay connection with rest of the supernodes. Any supernode looses its overlay connection only when there is no underlay (physical) connection to reach its successor overlay node. This means that the selected supernode is disconnected physically from the network and thus lacks a reachable path to its successor supernode. If such an overlay disconnection occurs in ad hoc networks, the communication undergoing in the overlay fails. A simulation based study on the criticality of unstable connections can be seen in our technical report [10]. In our simulation model, we calculate the ODR as shown in Eq. 4.

$$ODR = \frac{\text{Number of overlay disconnections in Simulation period } (OD_T)}{\text{Simulation Time}(T)} \quad (4)$$

B. Criteria of Supernode selection

We select supernodes based on different criteria. The criteria of selection of supernodes include the following criteria metrics:

- Randomly
- Based on number of 1-hop neighbors (node degree)
- Based on number of 2-hop neighbors
- Based on 2-hop CI index (Proposed)

For a given criteria, the specified criteria metrics are calculated for each node. For the evaluation of connectivity index, we take the overlay model as follows. For every network, we have N number of supernodes. For each instance of ad hoc network, C clusters are created. The clustering admission is based on the location of nodes. Each cluster has N/C serving supernodes. (Figure 4) In a given cluster, the nodes are sorted in the descending order of the criteria metrics for supernode selection. Then, the first N/C nodes in the sorted list are elected as the supernodes of the given cluster. Rest of the nodes associate with one of these serving supernodes. The supernodes within a cluster are cascaded randomly which further connect to supernodes in the subsequent clusters to form a Ring Overlay Topology.

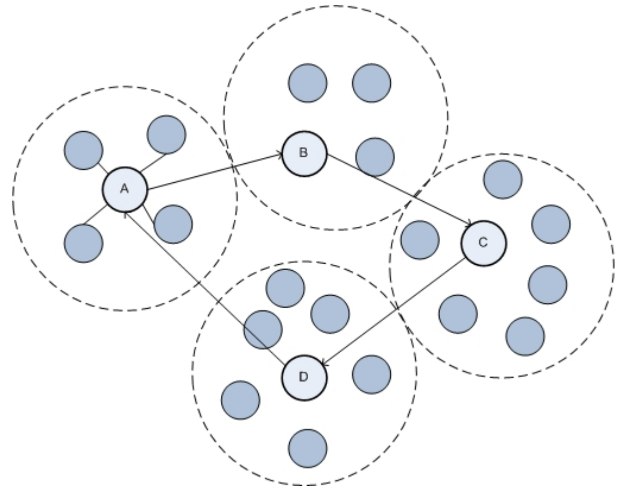

Fig. 4. Basic SN-based overlay scheme

C. Simulation Parameters

The input parameters of our simulation are summarized in Table I.

IV. Evaluation Scenario and Results

Here, we include results of simulation for the case of 150 nodes, uniformly distributed over an area of 2000 X $2000 sq.m$ with a radio coverage of $240m$.

A. With link Failure model

First, we apply link failure model with failure rate p. For each, p, we run 50 simulation cycles for each of 50 different instances of such topology in an attempt to average the result for different possible topologies that could exist with the given setting. Then we repeat the same for other values of p also. The results are plotted in Figure 5 and 6.We compare our proposed connectivity index with others.

In Fig. 5, we observe that the stability of the resulting overlay is very poor if the SNs are selected randomly (i.e. without considering the connectivity characteristics). The performance of such overlay with random SNs is outperformed by many fold when we make connectivity-aware overlay.The first observation thus leads us to an understanding that if such overlays are to be built in ad hoc networks with harsh conditions, SN elected without considering connectivity will result in unacceptable performance. Connectivity, in one or the other form, has to be a criteria of SN selection.

In Fig. 6, we observe the performance for 3 different metrics. At different link failure rates, the proposed CI has resulted in less average ODR, all the time, as compared to other metrics. Among the rest two, *size of 2-hop neighborhood* metrics has a slightly better performance as compared to *node-degree* metrics. The reason for the robustness of the proposed CI is attributed to the fact that the proposed CI is capable of reflecting the local connectedness of a node in a network more effectively. Due to this fact, a node with higher value of proposed CI is more connected and thus can withstand

TABLE I
SIMULATION PARAMETERS FOR CONNECTIVITY INDEX

Parameters of Topology	Value	Description
n	150	Number of nodes
R	240 m	Radio Coverage of each node
x1 $\times$ y1	2000 * 2000 m^2	Area in which nodes are distributed
dist	Uniform	Distribution of nodes in the specified area
Parameters of link/node failure model		
p		Probability of failure of each link/node
Statistical parameters		
Cycles	50	Number of times the same instance of topology is simulated
Instances	50	Number of graph instances that are simulated.

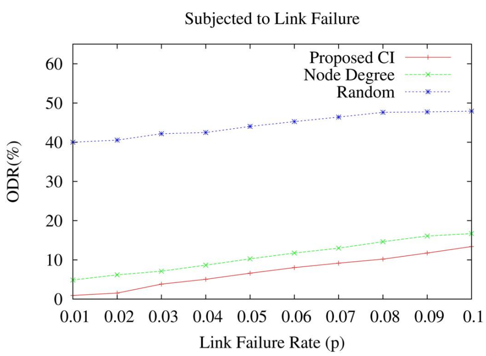

Fig. 5. ODR versus p - Link Failure

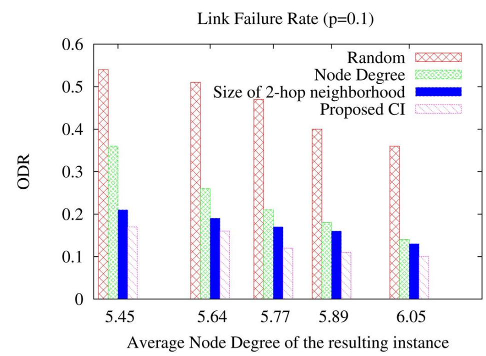

Fig. 7. ODR versus Network Instances (Average Node Degree) for $p = 0.1$ -Link Failure

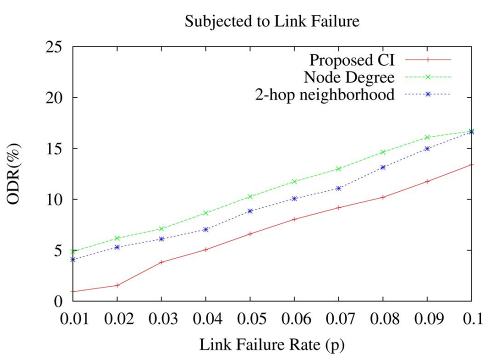

Fig. 6. ODR versus p (Contd)- Link Failure

more number of link failures than any other node with a CI anything less than this. Results showing the capability of the propsoed CI to integrate both of the other two metrics into it are discussed later and are demonstrated in Fig. 9.

The result clearly show that our proposed solution results in less number of overlay disconnections which implies that our connectivity index can be used to make better supernode selection strategies for forming robust overlays.

In Figure 7, we collect the statistics for different instances of the network graph for the same value of p (=0.1). Each instance of the graph is represented by the average node degree, which represents the connectedness of the resulting graph. We also see that our CI index based supernode admission is more effective when the network is less connected (in terms of average node degree). However, for networks with good level of overall connectivity, the effectiveness of supernode admission using CI decreases. This is mainly because when nodes in a network have,in an average, larger number of neighbors, more number of nodes will have a level of connectivity to withstand more number of random link failures. When more nodes have a good connectivity, even random election has a chance of selecting good SNs, and thus more robust overlays.

B. With Node Failure Model

We apply node failure model with rate p = 0.04. We collect statistics for different instance of the network graph (represented by the average node-degree). Figure 8 summarizes the result. Like in link failure model, the less connected a network graph is, the more effective is the connectivity based supernode admission.

Now, we illustrate the property of proposed CI that results in more robust overlay. In Fig. 9(a) and 9(b), we plot the relation

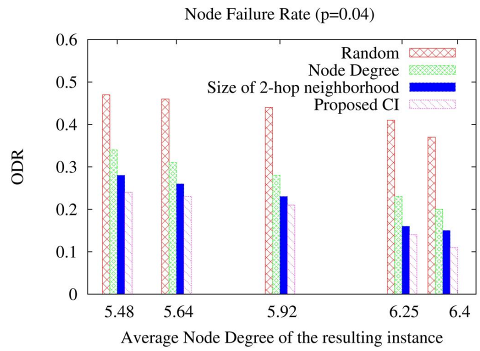

Fig. 8. ODR versus Network Instances (Average Node Degree) for $p = 0.04$ -Node Failure

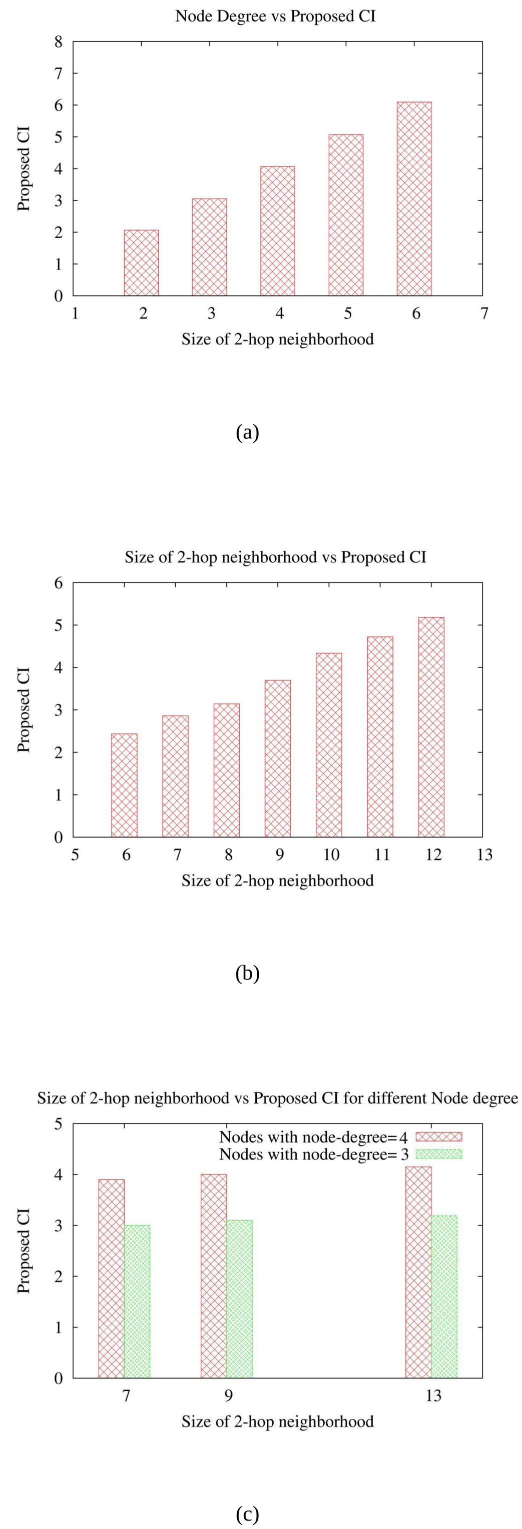

Fig. 9. Property of Proposed CI - its relationship with *node degree* and *size of 2-hop neighborhood* metric

between proposed CI and the existing connectivity criteria metrics. The results are the averaged simulation results for the network instances already defined in the simulation model. We see that the proposed CI of nodes increases with increasing *node-degree* metric as well as *size of 2-hop neighborhood* metrics. The proposed CI , thus, captures the property of both existing metrics individually.

Additionally, the proposed CI has the property of integrating both existing metrics. Refering to Fig 9(c), among the nodes with same value of *node degree* metric, nodes having bigger *size of 2-hop neighborhood* have higher CI value than those who have less number of 2-hop neighbors. Also, for the nodes with same number of 2-hop neighbors, the nodes with larger node-degree have, in average, higher value of CI. In general, the proposed index is capable to reflect both connectivity metrics: the *node-degree* metrics and the *size of 2-hop neighborhood* metric. In average, CI index can thus discriminate the connectivity of two nodes that have same number of 1-hop neighbors but different 2-hop neighbors and vice versa.

C. Complexity of Computation

Computation of node connectivity requires exchange of messages among nodes. A node has the information about the number of its neighbors. This is the minimum information maintained by each node. In this scenario, we comment on the complexity of the connectivity index.

Our 2-hop Connectivity Index (CI) requires the exchange of the node-degree information among the neighbors of the network. In this regard, irrespective of the size of the network, the connectivity calculation of each node requires message exchanges upto 2-hops. For any node v, the CI calculation requires node v to obtain the information of the edges corresponding to connection between nodes that are upto 2-hops away (Refer to Eq. 3).

In sparse ad hoc networks, the Connectivity Index calculations however are greatly simplified in terms of the complexity. The trick of the simplification lies on the fact that we view the 2-hop neighborhood of any node as a tree. Any mutual loops among 1-hop neighbors cause such nodes to reappear twice in the 2-hop neighborhood tree. With this assumption, we see an interesting property of proposed CI as follows.

If the two-hop neighborhood of a node is viewed as a tree, 2-hop criteria based CI (Proposed CI) of a node is equal to the sum of 1-hop criteria based CI of its neighbors.

If we consider the graph as a tree, we can easily show this. See our techical report [10] for proof.

This leads to a light weight implementation (in terms of messaging) to compute the 2-hop connectivity index. For example: Each node calculates its CI based in 1-hop connectivity criteria. And, then it can share this information (only the sum but not the individual information) to its one-hop neighbors. By doing this (i.e. by receiving the sum and not the individual component of the sum in Eq. 3 from second hop edges), it reduces the message exchange overhead geometrically. Also, the only messaging required is with one-hop neighbors. In any practical networks, neighbor nodes exchange informations in order to discover and maintain their neighbor-list and thus our CI calculation does not require any “beyond one-hop” message exchange. In practical network setting, this tree-assumption can lead to a simple implementation.

We had also compared by simulation, how the tree-assumption and the CI based on Eq. 3 differ. For the network setting specified in the simulation model, we found that the resulting SN selection is almost identical in most of the times. Refer to [10] for the details.

V. Conclusion

In this paper we proposed a “one-valued” node connectivity index (CI) for measuring the connectivity level of nodes in a given wireless ad hoc network. The proposed connectivity index was used as a criteria for selection of supernodes in the P2P overlay over ad hoc network. Conditioned to various rates of link-failure and node-failure, the robustness of the resulting overlay was compared with the robustness of overlays resulting from supernode selection based on either no connectivity criteria or with existing criteria. The results showed that the overlay using this criteria for selection of its supernodes is more robust in the terms of the overlay disconnections.

References

[1] S. A. Baset and H. G. Schulzrinne, “An analysis of the skype peer-to-peer internet telephony protocol,” *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pp. 1–11, April 2006.

[2] A. O’Driscoll, S. Rea, and D. Pesch, “Hierarchical clustering as an approach for supporting p2p sip sessions in ubiquitous environments,” *Mobile Wireless Communications Networks, 2007 9th IFIP International Conference on*, pp. 76–80, Sept. 2007.

[3] C. Bettstetter, “On the connectivity of ad hoc networks,” *The Computer Journal*, vol. 47, no. 4, pp. 432–447, Jul. 2004. [Online]. Available: http://www3.oup.co.uk/computer_journal/hdb/Volume_47/Issue_04/470432.sgm.abs.html;http://www3.oup.co.uk/computer_journal/hdb/Volume_47/Issue_04/pdf/470432.pdf

[4] Dousse and Thiran, “Connectivity vs capacity in dense ad hoc networks,” in *INFOCOM: The Conference on Computer Communications, joint conference of the IEEE Computer and Communications Societies*, 2004.

[5] T. Clausen and P. Jacquet, “Optimized Link State Routing Protocol (OLSR),” RFC 3626 (Experimental), Oct. 2003. [Online]. Available: http://www.ietf.org/rfc/rfc3626.txt

[6] S. Nelson, “A simulation study of connectivity metrics for wireless ad hoc networks,” *Thesis, Bucknell University*, May 2006.

[7] M.A.Rajan, M. Chandra, L. C. Reddy, and P. Hiremath, “A study of connectivity index of graph relevant to ad hoc networks,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 7, no. 11, pp. 198–203, Nov. 2007.

[8] M. Randic, “On characterization of molecular attributes,” *Journal of American Chemical Society*, 1975.

[9] ns 2, http://www.isi.edu/nsnam/ns/.

[10] J. Ghimire, M. Mani, N. Crespi, and T. Sanguankotchakorn, “A novel node connectivity index for wireless ad hoc networks,” 2009 ISSN: 0183-0570 , Ref: RR 09 011 RS2M.