CLUSTERED MOBILITY MODEL & EFFECT OF NODE CLUSTERING IN MULTI-HOP WIRELESS NETWORKS

VIKRAM SINGH RAWAT

Bachelor of Engineering in Electronics & Telecommunications

University of Pune, India

May, 2000

submitted in partial fulfillment of requirements for the degree

MASTER OF SCIENCE IN ELECTRICAL & COMPUTER ENGINEERING

at the

CLEVELAND STATE UNIVERSITY

December, 2004

This dissertation has been approved
for the Department of Electrical and Computer Engineering
and the College of Graduate Studies by

Dissertation Committee Chairperson, Dr. Chansu Yu

Department/Date

Dr. Vijay K. Konangi

Department/Date

Dr. Yongjian Fu

Department/Date

ACKNOWLEDGEMENTS

Many thanks go to my advisor, Dr. Chansu Yu, for making me do this thesis. I highly appreciate his advice and feedback which helped me in completing the thesis.

I would also like to thank Dr. Vijay Konangi and Dr. Yongjian Fu for their constant support and advice throughout the work.

I wish thanks to Lubo Song, Sridhar Kalubandi and Snehal Ganjigatti for their suggestions and opinions which helped me solve certain problems and added to my perspectives towards my work.

Special thanks to my family and friends for standing by me and boosting my morale during times when it seemed impossible to complete the thesis on time.

CLUSTERED MOBILITY MODEL & EFFECT OF NODE CLUSTERING IN MULTI HOP WIRELESS NETWORKS

VIKRAM SINGH RAWAT

ABSTRACT

An *ad hoc* network consists of wireless mobile nodes which can communicate without the aid of any existing network infrastructure. Mobility plays an important role in performance of ad-hoc networks and therefore mobility models used in simulations should represent realistic scenarios. Popular mobility models assume the initial node distribution and movement to be independently and randomly distributed over the simulation area. However in real life nodes usually tend to form clusters (concentrate in certain regions) leading to a non-homogenous node distribution.

The goal of this thesis is to study the performance of static wireless multi hop networks in presence of node clustering. For this purpose a mobility model that can generate clustered scenarios has been developed and implemented in Ns-2 and termed as the Clustered Mobility (CM) model. Extensive simulations were carried out using the CM model for static scenarios. Results from Random static scenarios have been used as the reference. It has been found that the performance of a static clustered scenario can vary significantly depending on the traffic pattern. In particular a scenario with larger proportion of intra cluster traffic performs better than a scenario with larger inter cluster traffic.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER I

INTRODUCTION

1.1 Background

Mobile Ad-hoc NETwork (*MANET)* is a collection of wireless mobile nodes configured to communicate amongst each other without the aid of an existing infrastructure. MANETS are *Multi-Hop* wireless networks since one node might not be in direct communication range of other node. In such cases the data from the original sender has to travel a number of hops (hop is one communication link) in order to reach the destination. The intermediate nodes act as routers and forward the data packets till the destination is reached.

Like other networks the performance of ad hoc networks is affected by its topology. However, in ad-hoc networks the role of topology becomes more critical due to mobility of the nodes. Simulation tools used for ad hoc networks studies employ mathematical models known as *mobility models* to generate various kinds of network topologies (static and mobile). It is important that Mobility models should be able to mimic the distribution and movement pattern of nodes in real-life scenarios. Often researchers employ indigenously designed models or popular mobility models (ex: Random Walk [3], Random WayPoint [10]) for performance comparison of various protocols. The concern with indigenously designed models is that they represent very specific scenarios not often found in real lives. Hence their use in ad hoc network studies is very limited. On the other hand popular models like Random Walk or Random Waypoint though simple and elegant, produce random movement patterns for each of the mobile nodes. However in real life the distribution and movement pattern of a *Mobile Node* (*MN*) is far from random and is influenced by various factors (time of day, traffic conditions, weather conditions etc.). Mobile Node (MN) is defined as a host or a router that changes its point of attachment from one subnet to the other.

1.2 Thesis Description

Most of the research on ad hoc networks is based on the assumption of random movement and distribution of nodes within the simulation area, which is true in some of the cases. However most of the network topologies around are not random but have a non-homogenous distribution of connections or nodes. The motivation for the thesis was derived from the fact that there has not been much work done based on non homogenous distribution and movement of nodes in ad hoc networks. Until now researchers have used indigenous methods to get non homogenous layout of nodes and there has been little work done towards developing a mobility model for such scenarios. Therefore, in this thesis a mobility model for clustered scenarios termed as the *Clustered Mobility* has been developed and used to study performance of static wireless multi-hop networks in presence of node clustering. This would help network designers in gaining in depth knowledge about such networks, and certain network design considerations arrived at, if necessary. An implementation of the model, "*realgen*", has been done in Ns-2 [16], the most popular Network Simulator package.

The model is based on the mechanism of preferential attachment [1] mechanism which states that, new node X has a higher probability of being attached to an existing node A as compared to another existing node B, if A enjoys larger number of connections compared to B. According to the CM model the whole simulation area is divided into smaller sub-sections also known as *sub-areas*. A simple algorithm based on the idea of preferential attachment (some sub-areas have higher probability of getting nodes based on current number of nodes) is then used to calculate the number of nodes in each sub-area. These nodes are then assigned random positions within the sub-areas respectively. A slight modification of the algorithm is used to choose destinations for the mobile nodes. The sub-areas with high density of nodes are the points of high interest and are known as *clustered sub-areas*. Even if nodes move, they have a higher tendency to move from one clustered sub-area to another. Example – In a shopping plaza, trade fair or rescue operation, it is highly unlikely that node distribution would be uniform.

Extensive simulations were carried out in Ns-2 and results from static random scenarios (using Random Waypoint model [10]) were used as the reference. Packet Delivery Ratio, Delay and Throughput were the metrics used to evaluate the performance as done in other previous work [8,12,34]. A definition of cluster, intra cluster and inter cluster traffic has been provided. A wide variation in performance for clustered scenarios was observed depending on the traffic pattern. In particular, performance improved as average hop count decreased. Also for a given percentage of intra cluster traffic, performance degrades with increased node clustering. In multi-hop networks certain nodes are unevenly burdened with the load of forwarding data packets compared to other nodes and can be classified as hotspots. These nodes form a bottleneck in the network. Location of hotspot nodes has been identified for both random and clustered scenarios. While the performance of a random scenario was used as a reference, it was observed that in such scenarios the performance is limited by the center of the simulation area. It has to be noted that, though the CM model is able to produce node movement the results presented in this thesis pertain to static wireless multi hop networks. Since Ad Hoc networks usually imply node mobility, in this thesis the term Static Wireless Multi Hop network has been used wherever necessary to emphasize the difference between the two.

1.3 Thesis Structure

Organization of the thesis is as follows: Chapter II gives an overview of the various network topology generation methods employed in ad hoc network studies. The chapter introduces various definitions and terms to the reader and then describes the goal of various topology generation methods, along with their characteristics, advantages and disadvantages. It gives a brief overview of various mobility modeling techniques and tools. In Chapter III the proposed Clustered Mobility model has been described in detail.

It describes the various related terms, properties and the algorithm used for the model. Chapter IV presents the simulation methodology and setup to study the impact of node clustering on ad hoc network performance. This chapter also presents the simulation results and supporting arguments. Chapter V describes the conclusions made from the obtained results and possible future work.

CHAPTER II

RELATED WORK

This chapter gives an overview of the various network topology or scenario generation methods employed in ad hoc network studies. The chapter begins by introducing the methods and terms generally used in topology generation for ad-hoc networks in Section 2.1. A brief description and goals of various mobility models, along with their characteristics, advantages and disadvantages have been discussed in Section 2.2. In particular the Random Waypoint which is the most heavily used mobility model is discussed in more detail than others in Section 2.3.

2.1 Introduction

Having a realistic network scenario plays an important role in experiments or simulation purposes to evaluate any communication network performance. While real-life experiments would provide the most accurate data, the difficulty to conduct one increases with number of nodes in the network. It becomes infeasible to conduct experiments for performance evaluation of large Ad-Hoc networks due to the cost and complexity involved. Thus usually simulation tools are employed to conduct the performance evaluation studies of such networks. In order to generate network topologies, generally 2 techniques are employed. The first one is to have real-life *traces* (trace is data of real-life scenarios). The second method is to use *synthetic models* (a synthetic model is mathematically representation of the characteristics of real life scenarios). The synthetic models developed for ad-hoc networks are known as *Mobility Models*. Synthetic mobility models can be further divided into *Individual Mobility* models and *Group Mobility* models. Though techniques to generate non homogenous or clustered scenarios can be considered another class of synthetic mobility models, lack of such models discourages us in doing so. Individual mobility models are also known as *Entity* mobility models [31]. As the name suggests, in entity mobility models the position and movement pattern of a *Mobile Node (MN)* is independent of other nodes in the simulation. But in group mobility models the movement pattern is dependent on one or more nodes. While a clustered scenario in some ways resembles a group based scenario, present group mobility models cannot fully characterize node clustering. Node clustering is characterized by a heavy tail node distribution.

Various individual and group mobility models have been developed for use in cellular or PCS (Personal Communication Systems) studies. Most of these models can be modified to adapt to the ad-hoc network simulation studies. However, only those mobility models that have been primarily developed for ad-hoc network simulations will be discussed in the sections below. We first describe the 'individual' mobility models in Section 2.2.1 and then the 'group' mobility models in Section 2.2.2. Thereafter, various techniques to generate clustered or non homogenous distributions have discussed in Section 2.2.3. An excellent description of various mobility models along with some of their interesting properties has been given in [31].

2.2 Description of Mobility Models for Ad Hoc Networks

2.2.1 Individual Mobility Models

- *Random Walk / Brownian Motion*: This model was prepared to model the erratic movement of various entities in nature. It has been reported in [3]. In this model each movement occurs for either constant time 't' or for constant distance 'd'. Speed and direction are chosen randomly between (V_{min} , V_{max}) and (0 , 2Π) respectively. After traveling for fixed time 't' or for distance 'd', MN chooses another speed and direction randomly and independently. Since the choice of next speed and direction doesn't depend on present or previous values, the process is memory less. When the MN reaches the boundary of simulation area it bounces off back into the simulation area with an angle determined by incoming direction and begins to move in this new direction. Disadvantage of this model is that it results in sharp turns and sudden stops. However this model is heavily used in simulations due to its simplicity.

- *Random WayPoint (RWP)*: This model has been proposed in [10] and overcomes the shortcomings of Random Walk model. The model is same as random walk with the exception of pause times. The model has been extensively used in ad-hoc network simulations and has become a reference model for other mobility models in simulation studies. For this purpose the model has been discussed in detail in Section 2.2.3.
- *Random Direction*: This model was developed to avoid the density waves (node clustering in a region), which occur in center of simulation area in the Random Waypoint mobility model [11]. In this model MN's select a random direction and speed and start moving in order to reach the simulation boundary, in that direction. Once boundary is reached the MN pauses for a specified time, then chooses another direction between (0 , Π) and the process continues. Since the MN's in order to pause have to travel to the simulation boundary, the disadvantage is that pausing at simulation boundaries increases the hop count and network partitions are more likely.
- *Boundless Simulation Area*: This model is not a memory-less model and hence a relationship exists between the previous direction, velocity and the current direction, velocity of a MN. Also unlike other models when a MN reaches the simulation boundary it reappears on the opposite side of the simulation area instead of bouncing back. This technique produces a torus shaped simulation area allowing MN's to travel unobstructed [36].

- *Gauss-Markov*: Model was originally proposed in [4] for PCS networks. Several implementations for the model exist. Initially each MN is assigned a speed and direction, thereafter at fixed intervals of time, *n*, the speed and direction of the MN are updated according to following formula

  $$s_n = \alpha s_{n-1} + (1-\alpha)\bar{s} + \sqrt{(1-\alpha^2)}\, s_x^{\,n-1}$$

  $$d_n = \alpha d_{n-1} + (1-\alpha)\bar{d} + \sqrt{(1-\alpha^2)}\, d_x^{\,n-1}$$

  where s_n and d_n are the time and speed at time interval *n*; α is the tuning parameter to vary the randomness and varies between 0 & 1; $\bar{s}$ and $\bar{d}$ are constants representing the mean value of speed and direction as $n \to \infty$; and $s_x^{\,n-1}$ and $d_x^{\,n-1}$ are random variables from a Gaussian distribution. The next location is calculated based upon current location, speed and direction of movement. Specifically at time interval *n*, an MN's position is given by the equations:

  $$x_n = x_{n-1} + s_{n-1}\cos d_{n-1}$$

  $$y_n = y_{n-1} + s_{n-1}\sin d_{n-1}$$

  where (x_n, y_n) and (x_{n-1}, y_{n-1}) are the *x* and *y* coordinates of the MN at the n^{th} and $(n-1)^{th}$ interval respectively and s_{n-1} and d_{n-1} are the speed and direction of the MN, respectively, at the $(n-1)^{th}$ interval. Since the current speed and direction values are based upon the previous speed or direction values, it is a memoryless process. This helps in eliminating sudden stops and sharp turns using the movement history of the MN to calculate future values.

- *Probabilistic Version of Random Walk*: As discussed in [31], this model describes the movement of a MN using a probability matrix, which is basically a state transition matrix. Each element of the probability matrix defines the probability with which MN can go from one state to another. There are three different states each for x and y positions of the MN. State 0 represents the MN's current (x,y) position, State 1 represents MN's previous (x,y) location and State 2 represents MN's next (x,y) location if MN continues to move in same direction. The probability matrix is shown below, where $\mathrm{P}(a,b)$ represents the probability of an MN going from state a to to state b.

$$\mathrm{P} = \begin{bmatrix} \mathrm{P}(0,0) & \mathrm{P}(0,1 & \mathrm{P}(0,2) \\ \mathrm{P}(1,0) & \mathrm{P}(1,1 & \mathrm{P}(1,2) \\ \mathrm{P}(2,0) & \mathrm{P}(2,1 & \mathrm{P}(2,2) \end{bmatrix}$$

  This model reduces the erratic movement associated with the random walk model by introducing probabilistic rather than random movement of MN's. However choosing appropriate values of $\mathrm{P}(a,b)$ for each scenarios may limit the use of this model unless traces are available.

- *City Section / Manhattan Mobility model*: This model is used to represent MN movement within a city section, which is basically a criss-cross of various streets and avenues. The streets and speed limits are based on city being simulated. Each MN starts at a fixed position on some street. It then randomly selects the destination position, which is also a point on some street. While deciding the movement path, the shortest path algorithm is used and safe driving characteristics (ex maximum speed limit and safe distance between two MN's) are kept in mind. This model provides a real-life scenario for a city, where people usually do not move in an independent and random direction and with constant speed, but according to some laws and restrictions. Various improvements which might include frequent pausing, acceleration, deceleration etc can be introduced to the model.

- *Obstacle mobility model*: In [2] an obstacle mobility model has been proposed, wherein obstacles (buildings etc) obstruct the movement path as well as wireless transmission. Obstacles are also treated as destination points and nodes move from one building to other. The pathways for node movement are modeled using Voronoi-diagrams. The selection of a pathway is done using the shortest path algorithm. The model is a step forward towards realistic mobility model, however, the selection of node destination and initial distribution is based on Random Waypoint.

2.2.2 Group Mobility Models

While a clustered scenario in some ways resembles a group based scenario, present group mobility models cannot fully characterize node clustering. Succeeding paragraphs describe some of the popular group mobility models and how they are insufficient to capture the characteristics of a clustered scenario.

- *Exponential Correlated Random Mobility Model*: This model uses a motion function to create MN movements [13]. The function used is

  $$b(t+1) = b(t)e^{-1/\tau} + (\sigma\sqrt{1-(e^{-1/\tau})^2})r$$

  where, *b(t+1)* is the next position and *b(t)* is the current position of MN; *r* is random Gaussian variable with variance σ; and τ adjusts the rate of change from MN's previous location to new location. However, selecting appropriate values of (σ,τ) for a given scenario is not easy.

- *Column Mobility Model*: According to this model [27], the MN's move around a line or column along the forward direction. A reference grid (line or column of reference points) is initially formed. The MN's are placed randomly around the reference points and are allowed to move around the reference point according to one of the entity mobility models (proposed is the Random Walk mobility model). The reference grid also moves according to a predefined distance and offset (between 0 to Π since only forward movement is allowed) and hence results in movement of the MN reference point. The new reference point is given as

  $$new\_reference\_point = old\_reference\_point + advance\_vector$$

  This model is useful in scanning and search purposes.

- *Nomadic Community Mobility Model*: This model [27] represents movement pattern of nodes that collectively move from one point to another. Each collection of MN's also called as a group has a logical reference point, unlike in column mobility model where each MN has a reference point. MN's in the group move around this logical reference point according to an entity mobility model. Therefore the nodes in nomadic mobility model enjoy more freedom of movement compared to the column mobility model.

- *Pursue Mobility Model*: This model [27] represents mobile nodes tracking a particular target. The new position of each mobile node is calculated as follows:

  $$new\_position = old\_position + acceleration(target - old\_postion) + random\_vector$$

where, acceleration (target – old_position) is information about object/MN being pursued and random_vector is the random offset for each MN. The value of random_vector can be obtained through an entity mobility model and should be limited to maintain effective tracking.

- *Random Point Group Mobility (RPGM) Model*: A popular mobility model to study group mobility scenarios in ad hoc networks known as Reference Point Group Mobility (RPGM) has been proposed by Hong et al in [34]. This is the most commonly used group mobility model. Since it is a more general model, the other three group mobility models (Column, Group, Pursue) can be implemented as specific cases of RPGM model. According to this model each group has a logical center and the movement pattern of logical center controls the member node movement pattern including speed, direction, acceleration, etc. Each group's logical center movement is defined by a Group Motion vector ($\overrightarrow{GM}$). $\overrightarrow{GM}$, can be random or predefined. Apart from that each member node has a reference point which in turn depends on the location of the logical group center. The member nodes move independently within a certain distance of their reference points. This vector which controls a node's independent movement is called random motion vector ($\overrightarrow{RM}$). $\overrightarrow{RM}$ has a length uniformly distributed within a certain radius centered at reference point and direction uniformly distributed between (0,2Π). Thus at any instant '*t*' the location of i^{th} node in j^{th} group is defined by following relation

$$X_{i,j}(t) = GM_j(t) + RM_{i,j}(t)$$

The MN's do not pause while the group is moving. The group nodes pause when the group reference point reaches the destination and all nodes pause for the same time.

Though this model can be used to simulate various scenarios, the definition of a group, restricts the motion of its members. A node in a group has a high correlation with other group member nodes. This might not be the case in clustered scenarios. For example in a trade fair some displayers attract lot of people. People coming to the show may come in groups or individually. Individual movement is usually independent.

- *Reference Velocity Group Mobility Model (RVGM)*: Proposed in [24], the authors claim/argue that similarity in member node movements is a more fundamental characteristic of mobility group rather than proximity in physical displacement. Thus each node is characterized by velocity $v = (v_x, v_y)^T$ where v_x and v_y velocities in x and y direction respectively. Each mobility group is characterized by a characteristic group velocity also called as *mean group velocity*. Mean group velocity acts as the reference velocity for nodes in a group. Specifically the node velocity is given by following relation

$$V_{j,i}(t) = W_j(t) + U_{j,i}(t)$$

where, $W_j(t)$ is a random variable drawn from Group velocity Distribution, $P_{j,t}(w)$, for group j; $U_{j,i}(t)$ is a random variable drawn from local velocity Distribution, $U_{j,t}(w)$, for member nodes of group j. Thus each group as a whole will follow a path which can be derived from any arbitrary distribution, $P_{j,t}(w)$. The member nodes within a group will have their own movement pattern around the reference point and can be derived from another arbitrary distribution $U_{j,i}(t)$. This model was prepared to see the effect of Network Partitioning in particular. Again in this model there is high correlation among the velocities of group members. Thus nodes within a group tend to have low relative velocities for longer time periods since they are derived from same group and node velocity distributions. This might not be the case in clusters, wherein a node might remain in the same clustered area for a while or leave it immediately to move to another area. Also using RVGM it is impossible to achieve scenarios where clustered areas are fixed but nodes move to and from or within the clustered areas, since for that purpose we need to have a zero reference velocity for all groups.

- *Grcmob*: In [23] 4 variants of group mobility models have been presented. A software module named *grcmob* has also been provided to generate the scenarios according to the 4 mobility models. The random waypoint model has been extended to incorporate the idea of group. The logical idea of group has been implemented using notion of sensitivity, which in-turn is has been characterized by 3 parameters which decide a node's group. The assumption is that each group has a fixed size and nodes are evenly distributed among groups. The authors of [23] have come up with conclusions regarding impact of group scenarios on MANET performance. Specifically effects of inter-group and intra-group communications and network partitioning have been highlighted which should be considered when designing networks where group formation is likely. The mobility models presented in [23] are not ideal to model clustered layout since node distribution among groups is considered to be even.
- *Virtual Track (VT)*: Virtual Track (VT) is a mobility model, which can model group mobility (including group dynamics like group split and merge) as well as individual and static scenarios [5]. The authors refer to the non-uniform, dynamic changing scenario described above, as "heterogeneous" group mobility scenario. The model consists of "switch stations" interconnected by "virtual tracks". The location of

switch stations can be user specified or random. Groups of nodes move on these virtual tracks towards a switch station. At the switch stations nodes can switch groups and groups can split or merge. Group membership is decided using a stability threshold value. The model can also represent individual and static nodes, which are again modeled using Random Waypoint model. The model seems to be specifically targeted for military scenarios rather than urban and sub-urban environment, where it is highly unlikely that people move in groups on roads or highways. Also the individual movement is not random.

2.2.3 Techniques for generating Non-Homogenous Topologies:

Only recently researchers have realized the shortcomings of random models and have started to focus on issues regarding non-homogenous node distribution and heterogeneous node movement behavior. While some authors have briefly mentioned techniques to generate non-homogenous distributions and particular solutions pertaining to them, others have developed mobility models to simulate such scenarios. The following paragraphs present a short description of such works.

In [32] transmission power control, clustering and connectivity issues for non-homogenous node distribution have been discussed; the authors haven't provided any method to generate such scenarios. [19] Discusses distributed power transmission management among the nodes to maintain strong connectivity. The authors have mentioned to employ a binary multiplicative branching process to construct simple clustered patterns. It is an iterative process: at first the total area is divided into 4 equal sub-areas. Two randomly chosen sub-areas get a fraction (1+β)/4 of the parent probability mass μ = 1, and the rest two get (1-β)/4. Here β is any constant less than one. In the second iteration each sub-area is again divided into 4 sub-areas and non-uniformly redistributes *its* probability mass into its four offspring sub-areas. Thus after j^{th} iteration the probability mass μ = 1 is non-uniformly distributed among 4^j sub-areas with area $1/4^j$, where $\begin{pmatrix} i \\ j \end{pmatrix} 2^j$ of these sub-areas $(0 \le i \ge j)$ come with probability mass $[(1+\beta)/4]^i[(1-\beta)/4]^{j-I}$. For each of the N nodes a uniform random number between 0 and 1 is generated and which according to the probability-mass-weighted ordering of the 4^j sub-areas correspond to exactly one sub-area. This results in a multi-fractal or clustered distribution of nodes.

A model for Static Clustered Layout has been proposed in [8]. Here the concept of power law distribution has been utilized. The simulation area is divided into smaller sub-areas. The Cumulative Distribution Function (CDF) of Bounded Pareto-Distribution, which is a simple power law, has been used to calculate the number of nodes in the sub-areas, which are then randomly distributed in the sub-areas. The CDF of Bounded Pareto distribution is given as follows

$$F(k) = \frac{1-(a/k)^{\alpha}}{1-(b/k)^{\alpha}}$$

The complexity in above model lies in selecting proper values for the parameters a, b, k and α, for a given scenario.

BonnMotion [15] is a mobility scenario generation and analysis tool implemented in Java. The tool can be used to create scenario files to be used in Ns-2 and GlomoSim/Qualnet [14] simulation packages. Apart from implementing the popular mobility models (RWP, Gauss-Markov model, Manhattan Grid model and RPGM), the tool provides two methods to generate non-homogenous node distributions. One method uses the concept of '*attraction points*'. One can specify the location of attraction points and their relative '*intensity levels*'. Nodes are placed within certain distances from the attraction points based on a Gaussian distribution whose parameters one can specify. An attraction point with an intensity *x* times higher than another point, would attract a node with a probability which is *x* times high. Another method distributes the simulation area into smaller areas with different node densities along its x- axis. Given the number 'n' of density levels, each of the n areas will contain a fraction of approximately,

$$\frac{2*k}{n*(n+1)},\ 1 \le k \le n,$$

of the nodes. (The density decreases from left to right). While this method is somewhat similar to the proposed clustered mobility model, the selection of optimal parameters (attraction points, intensity levels) is highly dependent on the user. Also since the intensity levels are fixed during the entire simulation, attraction points are fixed which is not true for various situations. Another mobility model, which is an extension of BonnMotion has also been reported in [26].

2.3 Random Waypoint Mobility Model

In ad hoc networks due to mobility of nodes the network topology varies continuously and thus the simulation results. Random WayPoint (RWP) mobility model was first used and described in [10]. Majority of the papers related to ad-hoc networks use this model in their studies. Specifically RWP has been used in [22, 30] for the evaluation and comparison of protocols. Due to its widespread use in studies of ad-hoc networks, some researchers [9, 11, 20, 33] have done a detailed analysis of the RWP model and have come up with its characteristic properties. It has been found out that the RWP model fails to provide a steady state distribution and care must be taken while using it in simulation studies, since the results and conclusions derived from such simulations can be misleading. Thus researchers using RWP in their studies should be aware of its characteristics and limitations. For this purpose, a description along with the various characteristics and shortcomings of RWP has been described in Section 2.3.1 & Section 2.3.2 respectively.

2.3.1 Description of RWP model

In general in RWP mobility model, the nodes are assigned initial positions randomly and uniformly. The same nodes may then choose to remain static or start moving. The node chooses a destination randomly and uniformly within the simulation area. The node speed is chosen uniformly from [V_{min}, V_{max}]. The node then moves towards the chosen destination with the chosen speed. After reaching the destination the node pauses or rests for a certain amount of time and then again starts moving towards a new destination. The whole process continues till the end of the simulation. While the pause time is fixed and node speed are chosen from uniform distributions, in principle they can be derived from any arbitrary distribution. The more general case being where some nodes are moving while others are static, also the pause time varies from node to node and can be derived from any arbitrary distribution. A specific case of above generalization is as follows – the simulation begins with all the nodes static for a constant pause time. After the pause time is over the nodes choose destinations randomly and uniformly within the simulation area. Also the nodes choose a speed with which to move towards the destination randomly uniformly within [V_{min}, V_{max}]. After reaching the destination the node again pauses for the constant pause time. While, V_{min} can be zero, in setdest (utility to generate scenarios in Ns-2) a positive non-zero value of V_{min} is chosen to avoid division by zero.

Different implementations of random waypoint exist. Within Ns-2, two different implementations of RWP exist as in '*setdest*' and '*mobgen*' modules. Within setdest the simulation begins with all nodes remaining static for the pause time. After that all the nodes start moving toward their destinations. After reaching their destinations nodes again pause and the whole process continues. But in mobgen the simulation begins with some nodes moving while the others remain static for the pause time. Afterwards some nodes may decide to pause or rest and while others may keep moving.

2.3.2 Properties / Weaknesses of RWP

Usually the simulations used to study performance of protocols in ad hoc networks are run for a fixed amount of time frame (900-1200 seconds) and employ a mobility model for scenario generation. Like various systems, simulations should reach a *steady state* also known as *stationary distribution* in probability literature [33]. Since results are usually represented as time-averages, reaching the steady state is crucial for simulations.

Apart from other drawbacks, it has been found out that the RWP fails to provide a steady state or stationary distribution of mobility within usual simulation time frames (900-1200 seconds). This means that mobility varies greatly during this time frame and hence the results derived from those. As a consequence the results are obtained by averaging over time do not represent the true picture or performance of any protocol. In succeeding paragraphs, different characteristics and drawbacks as found by various researchers, have been discussed.

In [33] authors have analyzed and characterized the random waypoint model and have come up with expressions for its stationary distributions. Specifically they have proved that PDF of initial *locations* and *speeds* of nodes differs from the distributions at later points of time. Also the PDF of both locations and speeds vary continuously over time and converge to a 'steady state' distribution known in probability literature as *stationary distribution*. They have derived the stationary distributions for speed, location and pause time for a node moving in rectangular area. If the initial speed and location (pause time if applicable) are sampled from the stationary distribution rather than the uniform distribution, convergence to stationary is immediate and no data need to be discarded. This saves simulation time and computational resources. An implementation of RWP using the stationary distributions has been provided as *mobgen-ss* module and is available at "http://toilers.mines.edu".

Yoon et al in [20] provide an analysis of the original RWP model and reasons why it is harmful in conducting simulations. Modifications have been suggested to improve the stability of the model. They have defined the *Instantaneous Average Speed* as,

$$v(t) = \sum_{i=1}^{N} v_i(t) / N$$

and have shown that it decreases as the simulation time progresses and eventually becomes zero. This is due to the fact that nodes with low speeds (ex 0.1 m/s) that have to travel large distances (1000m) , may never complete their travel within the simulation time (900 sec) and are stuck/trapped with such low speeds. Such low speeds influence the average speed. Time average of the performance metrics cannot be measured by varying the maximum speed; since taking the time average is based on the expectation that the mobility model reaches a steady state and the average node speed is $V_{max}/2$, however it doesn't happen in present implementations of RWP. They have provided with a simple modification to the RWP (having a minimum positive speed), which aids in the average node speed and thus the performance metrics, in faster stabilization to their ideal stationary values.

In [9] it has been proposed that in RWP the structure of the node distribution (asymptotic stationary distribution) is the weighted sum of three independent components: the static, pause and mobility component and their effects on the stationary distribution.

$$f_x(x) = f_s(x) + f_p(x) + f_m(x)$$

The node distribution $f_x(x)$ is composed of 3 distinct components, the static, the pause and the mobility component; where, $f_s(x)$, $f_p(x)$ and $f_m(x)$ represent the respective likelihood functions. Accurate approximations for them have been derived. $f_s(x)$ accounts for the fact that node can remain static for entire network operation time. Pause component $f_p(x)$ accounts for time that a mobile node "rests" before starting a new movement period. Mobility component $f_m(x)$ accounts for time that a mobile node is actually moving. The influence of velocity 'v' on $f_s(x)$ is less evident. In general higher 'v' causes smaller movement periods and hence a more "uniform distribution". At extreme values of pause times velocity has negligible effect on the node distribution. Small values of pause times result in significant border effect and thus non-uniform node distribution, whereas large pause times result in uniform node distribution. A method to achieve the steady state fast has also been described. This helps in reducing the computational resources and hence the simulation time drastically.

A brief survey of the RWP model has been done in [31]. The average MN neighbor percentage for MNs using the RWP model has been plotted and it shows that the there is a high variability in the average neighbor percentage for some initial period into the simulation, which would produce variability in the simulation results until the simulations are run for longer periods of time. They have proposed 3 methods to overcome such initialization problems. The first is to discard initial 1000 seconds of simulation data. The second is to save in files, the positions of nodes after a period long enough to pass the high variability stage. Then use these files as the starting position in all other simulations. The third is to use the initial positions and speeds derived from distributions representative of the stable state.

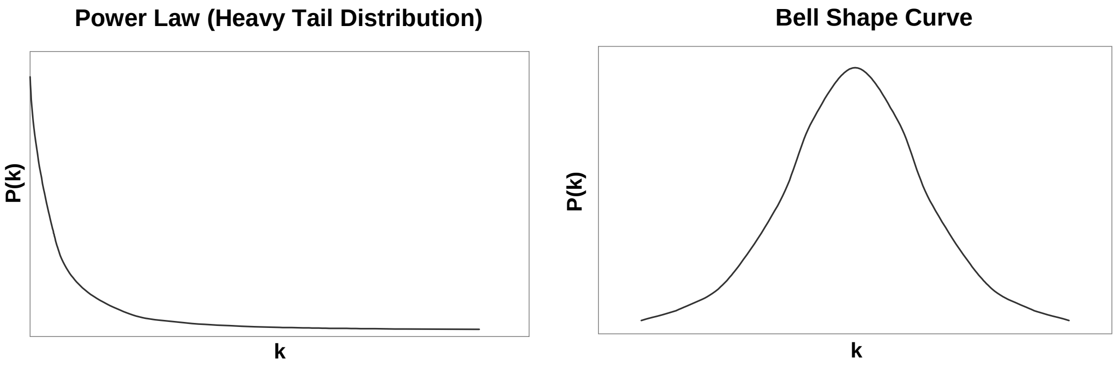
CHAPTER III

THE CLUSTERED MOBILITY MODEL

This chapter gives a detailed explanation of the proposed mobility model, which can be used to create clustered scenarios for ad-hoc network simulations. The basic terms and definitions used in developing the model are described in Section 3.1. It also gives a brief overview of the mechanics of a mobility model in general. The logical description of the model is described in Sections 3.2. Section 3.3 describes the properties of the proposed model for static scenarios. Section 3.4 presents the algorithm used to implement the model. Important factors that should be kept in mind while generating scenarios using the model are also presented wherever suitable.

3.1 Introduction

For many years most of the studies related to networks have assumed the structure/layout of the network (also known as *network topology*) to be *Random*. The plot of the fraction of nodes having *k* connections follows a bell curve as shown in Figure 1(b) and is a characteristic of random networks. This implies that the number of nodes having connections that deviate from the average value significantly is very low. In addition, all the nodes in random network have roughly the same number of connections to other nodes in the network.



(a) (b)

Figure 1. Connection Distribution among nodes

However, recently it has been found out that majority of such complex networks are not random; instead there are some nodes which enjoy high number of connections. These nodes are called as *hubs*. Hubs account for almost 80% of the network connections whereas the number of these hubs is very few, approximately 20%. Such networks are also called as *scale-free* networks [1]. Hubs can have hundreds, thousands or even millions of links, in this sense the network appears to have no scale. *Power laws* represent systems in which a few nodes called hubs, represented by the long heavy tail of the curve, dominate. Power laws, Figure 1(a), unlike bell shape curves do not have a peak; instead they are continuously decreasing function. Prior to the above-mentioned discovery, in order to generate random network topology, the nodes were picked randomly and connections made to them in an arbitrary manner. But to mimic the network topology of scale-free networks, 2 basic mechanisms namely *Growth* and *Preferential Attachment* have been presented in [1]. Growth refers to the increase in number of nodes as time progresses. Also a new node is extremely likely to attach to an existing node, which has a larger number of connections than the other and is called as preferential attachment process.

In almost all mobility models before the simulation can begin each node is assigned a position within the simulation area according to some criteria. This initial placement of nodes can also be called as the *initial layout*. Once the simulation starts the mobility model decides the movement pattern of nodes. Thus the purpose of a mobility mode is two fold, one to assign the initial layout and second to generate the movement pattern. Assuming the whole simulation area to be divided into equally sized smaller sub-areas, one can calculate the instantaneous number of nodes in each sub-area also called the *sub-area population*. A frequency plot of the sub-area population is termed as *spatial distribution* of nodes.

Since the RWP mobility model chooses the initial positions of nodes independently and randomly, the spatial distribution of nodes for the initial layout is a Bell shape distribution. The initial layout in which the nodes are randomly and independently distributed is termed as *random layout*. The RWP model has been implemented in Ns-2 and the modules are known as 'setdest' and 'mobgen'. The usage of setdest is as follows

setdest -n -p -s -t -x -y

where, n = total number of nodes; p = maximum pause time; s = maximum speed; t = total simulation period; x, y are the x, y co-ordinates respectively of the simulation area. A static scenario can be generated using RWP if the pause time is selected to be same as the simulation time.

However, the mobility model presented in this thesis utilizes the concept of preferential attachment thus leading to a heavy tail spatial distribution of nodes. Clustered sub-areas in our mobility model are analogous to hubs. They dominate the simulation area in the sense that they have a large fraction of total nodes. The CM model was implemented by modifying the 'setdest' module and has a similar usage.

3.2 Logical Description

Since our model extends the RWP mobility model, the method of selecting pause times and node speed are the same. The difference lies in selecting the initial layout and waypoints. The whole model can be visualized to consist of two steps, the first being to generate an initial layout and the second being the selection of a waypoint/destination to induce mobility. For static scenarios, only the first step is executed. The entire simulation area is logically divided into even number of smaller sub-areas numbered sequentially starting from 0, as shown in Figure 2. All the sub-areas have equal dimensions. The

12	13	14	15
8	9	10	11
4	5	6	7
0	1	2	3

Figure 2. Division of Simulation Area

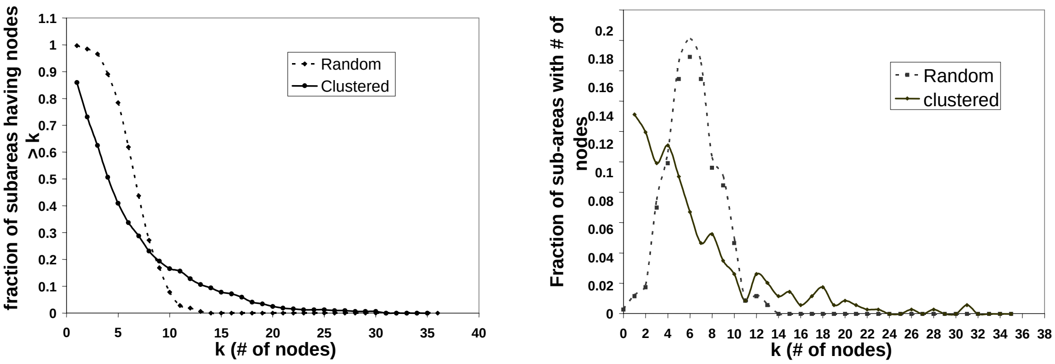
number of sub-areas can be user-specified. Each sub-area is then assigned a minimum number of nodes. The purpose of assigning a minimum equal number of nodes to each sub-area is two fold, one to prevent network partitioning and the other to provide an initial equal ‘parameter/popularity’ value to each sub-area. In our model, the ratio of present number of nodes in a sub-area to that of present number of nodes in total simulation area is the *parameter* chosen to quantify the popularity effect.

To generate *initial layout*, we first calculate the number of nodes in each sub-area also called the *sub-area population* using an algorithm based on Preferential Attachment. Since, initially all sub-areas have an equal number of minimum nodes, all sub-areas have equal parameter value and the probability of a sub-area being assigned the next node is equal. But as a new node is assigned to a sub-area its parameter value increases proportionally. After a few other nodes have being assigned a sub-area, some sub-areas will have larger parameter values than others and will be the preferred areas. It is important to note that in our model, growth refers to the process of assigning nodes one by one to the logical sub-areas. During the growth process the sub-area population changes and thus its corresponding parameter value. The growth process ends when all the nodes have been assigned a sub-area. This population is then used to calculate the corresponding *final parameter* or popularity values. Once the sub-area population is known, the nodes are randomly assigned a position within the sub-area boundary. A random number generator is used to choose among sub-areas based on the weighted parameter values. The use of a random number generator ensures the clustered sub-areas to be randomly distributed within the simulation area.

To *induce mobility*, waypoints are selected based on the fixed final parameter values. Again a random number is generated and based on the *final* parameter values a sub-area is chosen. Within the chosen sub-area boundary the destination is chosen randomly. The node selects a speed, which is uniformly distributed between (0-V_{max}) and starts moving towards it. Once the node reaches the destination, it waits for the pause time. After the pause time is over a new waypoint is selected again and the whole process continues till the end of the simulation period.

3.3 Properties of the Model

Figure 3 shows two kinds of node distribution for 100 nodes for both random waypoint and clustered mobility model. In Figure 3(a) node density distribution for random layout decays rapidly as k increases, while for clustered layout it decays slowly.

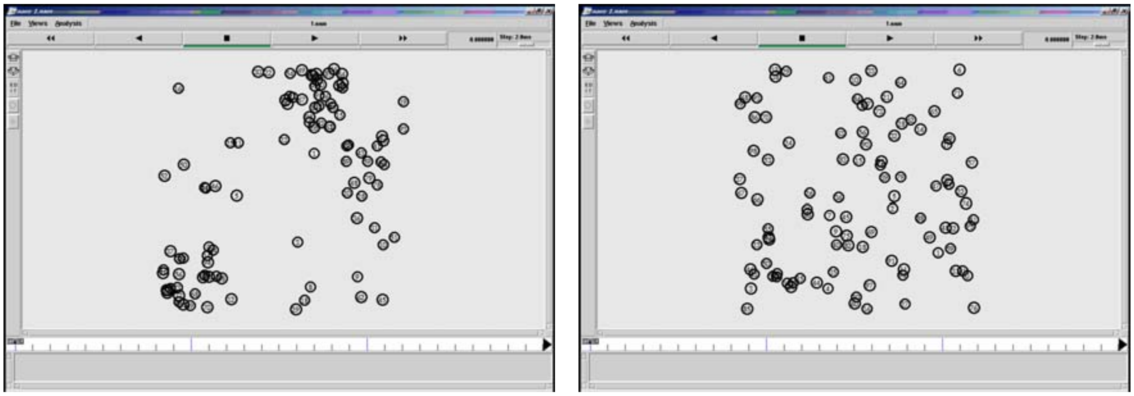


(a) 1-CDF (b) PMF/Frequency

Figure 3. Node Density distribution (Static Layout)

Figure 3(b) shows the plot of the frequency of sub-area population where the Random scenario has a bell shape whereas the clustered scenario has a shape of a power law.

A *nam* shot of both clustered and random scenario is also shown in Figure 4(a) and Figure 4(b) respectively. Nam is the animation tool used in Ns-2 and the snapshot of the scenario using it has been called as nam shot.



(a) Clustered Scenario (b) Random Scenario

Figure 4. Nam Shots

In this thesis a *Cluster* is defined to be a sub-area whose population equals or exceeds a threshold number of nodes. The threshold has been calculated according to the following method: A node distribution graph for random layout of 100 nodes is first plotted as shown in Figure 3. The number of nodes corresponding to the fraction representing the 0.05 or 5% fraction of sub-areas (on the right hand side of mean) is chosen as the threshold. This definition yielded a value of 10 nodes, for scenarios consisting of 100 nodes. With such a definition in random scenarios, approximately less than 5% of the sub-areas have nodes greater than the threshold i.e. clusters are rarely present in random scenarios. However in cluster scenarios approximately 20% of the sub-areas are Clusters (>= 10 nodes). It has also been found out that these 20% sub-areas contain more than 75% of the node population. With this fact one can assume that clusters contain majority (greater than 75%) of the nodes and thus majority of the traffic connections would originate or terminate in clusters. In other words the traffic pattern can be thought of to consist of inter cluster and intra cluster traffic connection pairs. *Intra Cluster* traffic is defined as the connection pairs whose source and destination lie within the cluster. For our scenario (1000x1000m^2), intra cluster traffic would be most of the times one-hop communication. Similarly *Inter Cluster* traffic is defined as the connection pairs whose source and destination lie in different clusters. Inter cluster traffic would be most of the times more than one-hop communication.

Associated with each sub-area is a parameter, which signifies its popularity. This parameter is a deciding factor in selecting the initial positions and destinations. Higher the parameter value, higher is the probability of a node being assigned the corresponding sub-area as destination or initial position. The Preferential Attachment is a linear process

[1]; therefore the parameter was chosen to be the ratio of number of nodes presently in sub-area to the total number of nodes presently in the simulation area. Whenever a sub-area is assigned a node, its probability increases by a factor of 1/Present number of nodes in system, compared to other sub-areas. To give control over the *Degree of Clustering* there are a variety of methods. One way is to add a constant ‘c’ to the parameter value of a sub-area every time a node is assigned to that sub-area. Assigning a non zero value to 'c' increases the Probability by a factor of '1/Present number of nodes in system + c' and results in a non-linear growth process. A large value of 'c' implies higher node clustering and usually results in only one or two sub-areas having all the nodes. Another method to increase the degree of clustering is to square the number of nodes, while calculating the parameter values.

3.4 Algorithm

Algorithm for initial layout

```
1-  Divide the simulation area into even number of sub-areas
2-  Assign minimum number (default 1) of nodes to sub-areas
3-  For remaining number of nodes perform the following steps
        I - for (i=0; i<max_sub_areas; i++)
            P[i] = (Present # of nodes in sub-area[i] / Present # of nodes in total simulation area)
        II- CDF[sub_area(0)] = P[sub_area(0)]
        III- for (i=1; i<max_sub_areas; i++)
                CDF[sub_area(i)] = CDF[sub_area(i)] + P[sub_area(i+1)]
        IV- for (i=0; i<max_sub_areas; i++)
                rand_no  =  random number between 0 and CDF[max_sub_areas]
                if (rand_ no  > CDF[sub_area(i)] ) && (rand_no <= CDF[sub_area(i+1)] )
```

```
                choice =  i
        V - Increment the number of nodes in the chosen sub_area by 1
        VI- Decrement Total nodes by 1
        VII - Goto step (3)
5-  Distribute the assigned number of nodes in a randomly within respective sub-area.
```

Algorithm for Choosing the Waypoint

```
1-  Repeat steps 4 to 9 till the end of simulation period
2-  for (i=0; i<max_sub_areas; i++)
                P[i] =  # of nodes in sub-area[i] / Total Nodes
3-  Calculate  CDF[sub_area 0] = P[sub_area 0]
        for (i=1; i<max_sub_areas; i++)
                CDF[sub_area(i)] = CDF[sub_area(i)] + P[sub_area(i+1)]
4-  Rand_no  =  random number between 0 and CDF[max_sub_areas]
        if (rand_ no  > CDF[sub_area(i)] ) && (rand_no <= CDF[sub_area(i+1)] )
        choice =  i
5- Select Destination = Random position within sub_area[choice]
6- Start moving towards the destination with a random speed (uniformly chosen from 0-Vmax)
7- After reaching the destination, pause for the pause time P(time)
8- After the pause time is over go to 4
```

CHAPTER IV

SIMULATION RESULTS & DISCUSSION

This chapter describes the simulation setup to study the impact of node clustering on static multi-hop wireless network performance and the results obtained from the simulations. Packet Delivery Ratio, Average Delay and Throughput are the 3 performance metrics used in this thesis. Their definition is given below:

- Packet Delivery Ratio – It is defined as the ratio of total data packets received at the destinations to those generated by the sources.
- End-to-End Delay – This is defined as the time required for a packet to travel from source to destination.
- Throughput – It is defined as the amount of data successfully delivered from the source to the destination in a given period of time.

Though, the proposed Clustered Mobility model is able to generate movement patterns, at present only *static* scenarios have been evaluated. Section 4.1 provides the details of the simulation setup. Section 4.2 presents the results and discussion.

4.1 Simulation Setup

This section discusses the simulation methodology and setup details to assess the impact of clustered scenarios on performance of static wireless multi hop networks. Clustered scenarios were generated using "*realgen",* an implementation of the Clustered Mobility model. It is an extension of the *setdest* module found in Ns-2 and has a similar usage. Random scenarios were generated using setdest module, which implements the Random Waypoint mobility model.

All the simulations were carefully conducted using the *Ns-2* (version 2.27) network simulator. All the simulations consisted of 100 nodes. Static scenarios were generated by making the pause time equal to simulation time. Simulation area of size 1000 x 1000 m^2 was selected. A radio transmission range of 250 meters and a two-way ground propagation model with a channel bandwidth of 2 Mbps was chosen for all simulations. The simulation area was logically divided into 16 sub-areas, all of equal size and shape. Thus a 1000 x 1000 m^2 simulation area results in a sub-area of size 250 x 250 m^2. 802.11 MAC protocol with RTS CTS is used for medium access. AODV (Ad hoc On-demand Distance Vector) routing protocol [7] is used to maintain and find routes.

All the scenarios, for both random and clustered mobility models, had all the nodes reachable so that there is no network partition in any of the cases. This was done in order to study the impact of node clustering rather than network partitioning. Though network partitioning is one of the consequences of node clustering (since it results in dense sub-areas but overall sparse distribution) and can result in severe performance degradation for certain cases, care was taken to avoid the effect of network partitioning on simulation results.

CBR (constant bit rate) traffic files with 46 sources and 74 connections were generated using the *cbrgen* utility in Ns-2. Packet rates employed were 0.5, 0.8, 1.0, 1.2, 1.4 and 2.0 packets per second. The packet size was 512 bytes. Since the performance results can vary from scenario to scenario, to get average results a number (8) of simulation runs were performed by varying the scenario files for both random and clustered layout.

4.2 Simulation Results & Discussion

4.1.1 Packet Delivery Ratio

In general, as the percentage of inter cluster traffic increases (or percentage of intra cluster traffic reduces), the performance of clustered layout degrades. As discussed in Section 3.3 intra cluster traffic can be assumed to be one hop whereas inter cluster traffic consists of two or more hops. The intra-cluster traffic can be delivered with less overhead and the increase in performance with increasing proportion of intra cluster traffic can be attributed to this reason. Figure 5 depicts the PDR for varying percentages of intra-cluster traffic for a clustered scenario for 1000 x 1000 m^2. The PDR for random scenario has also been shown as a reference.

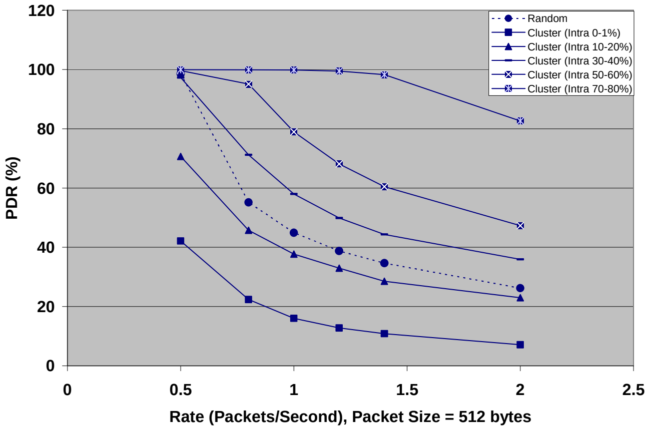


Figure 5. Packet Delivery Ratio, Area = 1000 x 1000 m^2

As can be seen from the graph the performance of a clustered layout varied significantly depending on the traffic pattern. The difference in performance could be as much as 87% (for intra = 70-80% & intra = 0-1% at a packet rate = 1.4 packets/sec) depending on traffic load and pattern. Since the connection pattern (source destination pairs) are same for varying packets rates, for a particular scenario, the hop count for all packet rates would be the same. With this fact in mind, hop counts were calculated for varying percentages of intra cluster traffic.

TABLE I: AVERAGE HOP COUNT & INTRA CLUSTER TRAFFIC PERCENTAGES

Scenario	**Average Hop Count**	**One hop connection OR Intra Cluster Traffic (%)**
Cluster	1.25	70 – 80
Cluster	2.0	50 – 60
Cluster	2.75	30 – 40
Cluster	4.0	10 – 20
Cluster	5.0	0 – 1
Random	3.0	10 – 20

TABLE I presents the intra cluster traffic percentages and corresponding average hop counts. As the percentage of intra cluster traffic increases the average hop count decreases.

4.1.2 End to End Delay

Figure 6 shows the end-to-end delay for varying percentages of intra cluster traffic in clustered scenarios for 1000 x 1000 m^2.

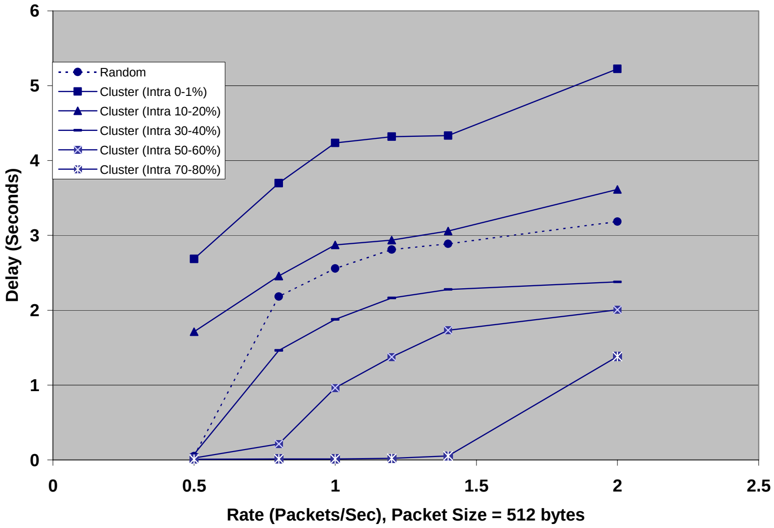


Figure 6. End-to-End Delay, Area = 1000 x 1000 m^2

The graph suggests that as the intra cluster traffic increases the delay decreases. This is again due to the fact that larger intra cluster traffic results in lesser average hop count. As the average hop count or path length increases data packets suffer longer delays at intermediate nodes. The difference in delay could be as much as approximately 4.5

seconds (for intra = 70-80% & intra = 0-1% at a packet rate = 1.4 packets/sec) depending on traffic load and pattern.

4.1.3 Throughput

It has been shown in [28] that for random wireless multi hop networks the maximum end to end throughput available to each node is proportional to ($1/\sqrt{n}$), where *n* is the number of nodes. Thus as the number of nodes approaches infinity the capacity approaches to zero. However, the authors in [21] argue that apart from the number of nodes in a network the traffic pattern also plays an important role in determining the available capacity to each node. Assuming a uniform node density the capacity available to each node is bounded by the following inequality

$$\lambda < \frac{C/n}{L/r}$$

where, C is the total one hop available capacity, n is number of nodes, L is the expected path length available from source to destination and r is the fixed radio transmission range. The above equality shows that as expected path length increases, the bandwidth available to each node decreases. Therefore, the traffic pattern has a great impact on scalability. As can be seen from the graph in Figure 7 as intra cluster traffic percentage increases (i.e. average path length or hop count decreases) the throughput also increases.

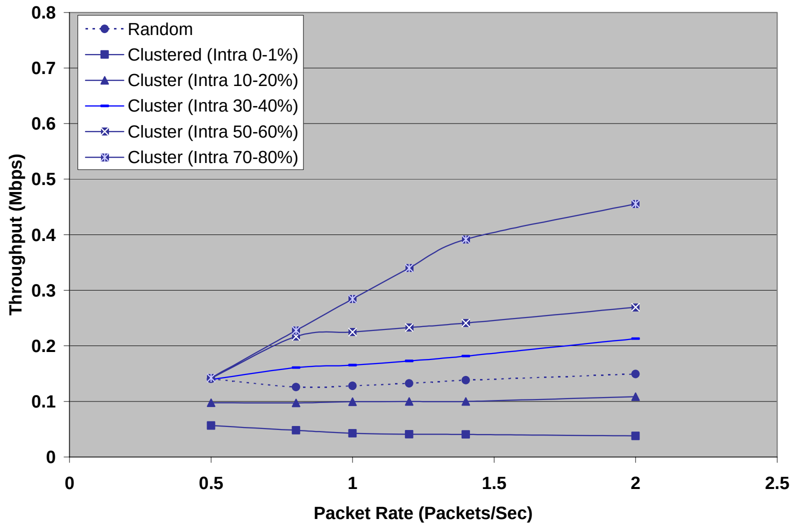


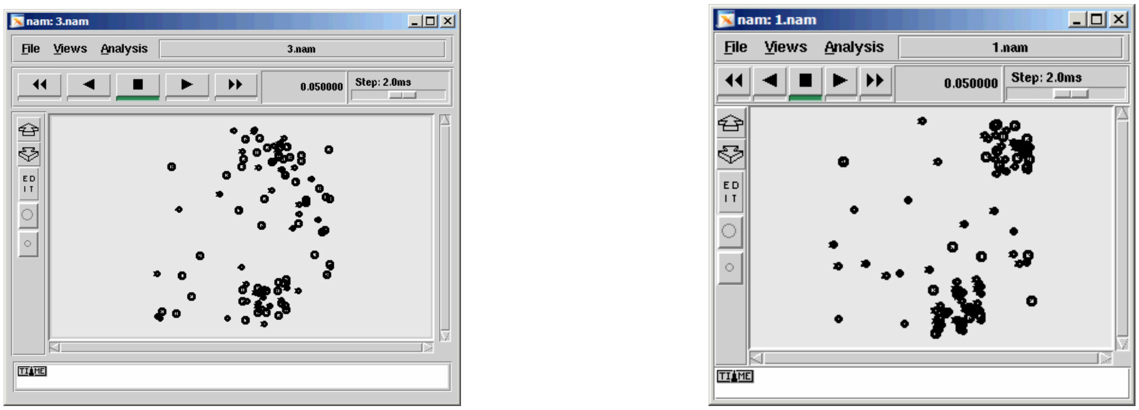
Figure 7. Throughput, Area = 1000 x 1000 m^2

The throughput usually increases till a certain traffic load after which it saturates. The knee of the throughput curve can be assumed to be the maximum throughput available to the network for that scenario. The maximum available throughput for various percentages of intra cluster traffic is seen to be proportional to the path length or hop count with only minor deviations.

4.1.4 Other Results & Observations

Increased Node Clustering degrades Performance

Using one of the methods as described in Section 3.3 scenarios with higher degree of clustering was generated. For comparison purposes the sub-area populations and Nam shots of two clustered scenarios with different degree of clustering are provided in TABLE II and Figure 8 respectively.



(a) Lower Node Clustering (b) Higher Node Clustering

Figure 8. Nam Shots for 2 different degrees of Node Clustering

TABLE II: SUB-AREA POPULATIONS FOR 2 DIFFERENT DEGREES OF NODE CLUSTERING

Sub-Area Number	**Number of Nodes**	
	Lower Node Clustering	Higher Node Clustering
0	24	1
1	6	1
2	11	42
3	1	1
4	10	3
5	1	3
6	1	3
7	1	5
8	7	1
9	2	1
10	1	1
11	1	1
12	1	1
13	17	1
14	1	1
15	15	34

A higher degree of clustering means fewer clusters with larger effective node density. With approximately the same percentage of intra cluster traffic, the performance of a clustered scenario depends on the degree of clustering: a scenario with lesser degree of node clustering performs better than a scenario with relatively higher degree of node clustering. This is intuitive since for same path length or average hop count, higher node density will result in higher contention for available bandwidth.

As can be seen with higher degree of clustering there are only 2 clusters as compared to 4 with a lower degree of clustering. This means that the node density ratio of the two scenarios is approximately two times. This difference is reflected in the capacity or throughput of the network which gets scaled by a factor of 1.4 or $\sqrt{2}$ times. This in turn affects the PDR as shown in Figure 9.

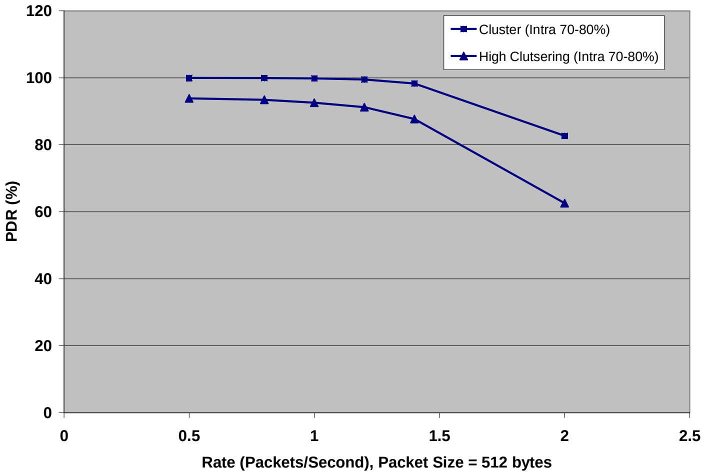


Figure 9. PDR Comparison with different degree of Node Clustering

Lesser degree of node clustering results in lower delay compared to higher degree of node clustering as shown in Figure 10. The difference becomes significant at higher packet rates, when the offered packet load exceeds the available bandwidth.

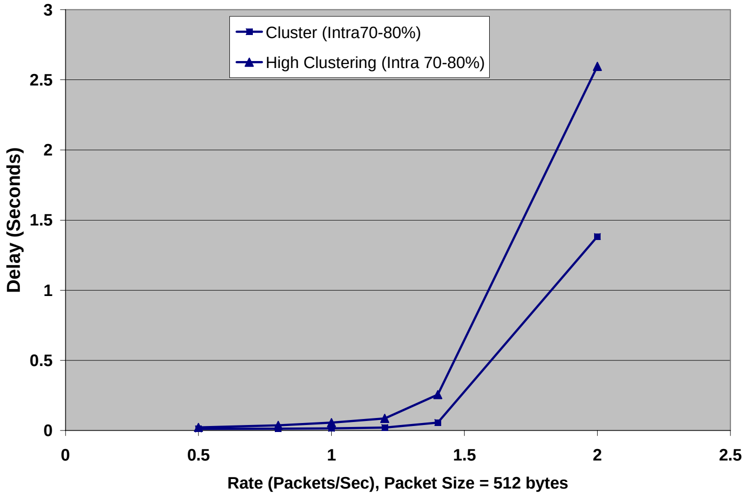


Figure 10. Delay Comparison for different Degree of Clustering

Also as node density increases it results in lower throughput (Figure 11) since the nodes have to share the constant bandwidth. As mentioned above the node density is almost twice in case of higher node clustering. This results in a capacity ratio of approximately 1.4 or $\sqrt{2}$ according to the results presented in [28].

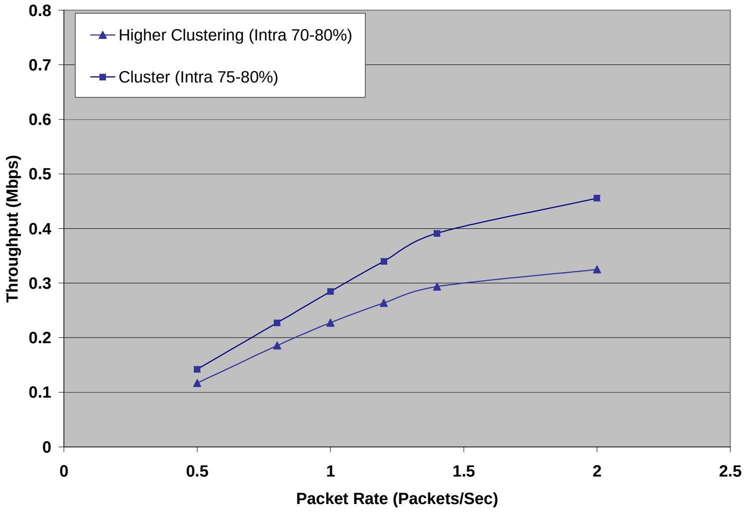
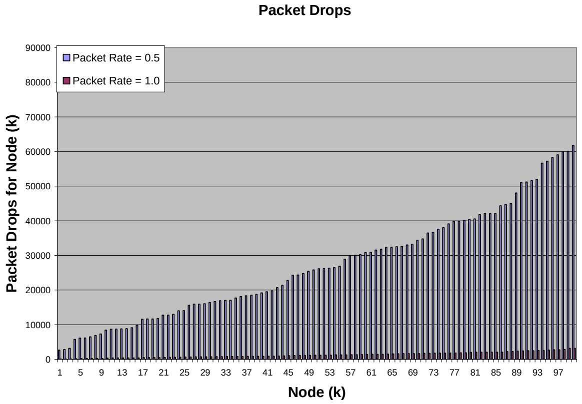


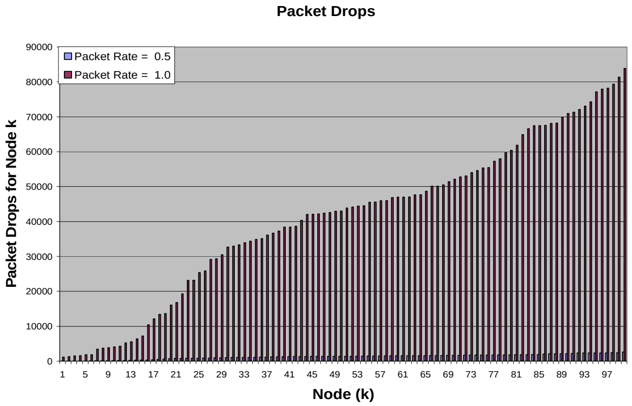
Figure 11. Throughput Comparison for 2 Different Node Clustering

Performance Degradation due to Existence of Hotspots

The performance in terms of PDR, Delay and Throughput of both random and clustered scenarios degrades with increasing packet rates as shown in Figure 5, Figure 6 & Figure 7 . On analyzing trace files for the dropped packets for 2 different packet rates (0.5 & 1.0) it was found that certain nodes drop significantly more packets than others. The total packet drops in ascending order for each node in the 2 scenarios are shown in form of bar graphs in Figure 12.



(a) Random Scenario



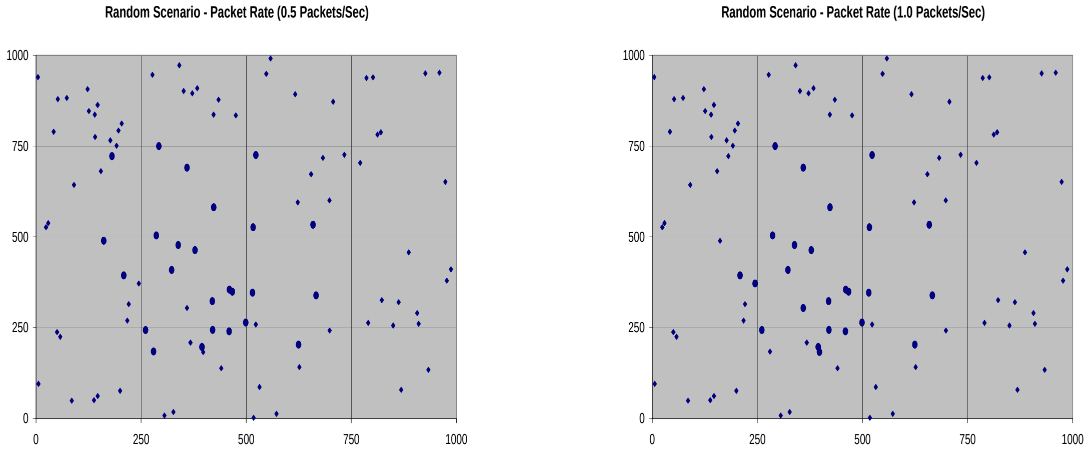
(b) Cluster Scenario (Intra 30-40%)

Figure 12. Total packet drops per node (Ascending Order)

For both packet rates the trend in packet drops is same, except that higher packet rates result in larger packet drops. The main cause for the degrading performance (both random and clustered scenarios) is the existence of certain bottleneck or hotspot nodes. The reason for formation of hotspot nodes is that, due to source and destination nodes far apart, data packets need forwarding. Certain nodes get unevenly loaded due to this forwarding traffic, depending on the routes chosen by routing protocol. These nodes drop more packets as compared to others due to high congestion and contention.

Hotspot Locations in Random Scenarios

The top 25% of nodes with the highest number of packet drops in each scenario were identified and were found to account for approximately 40-45% of total packet drops. These nodes can be classified as the *hotspot* nodes forming a major bottleneck in network performance. Location of hotspots for random scenarios in a 1000 x 1000 m^2 simulation area, are shown below in Figure 13.



(a) Packet Rate = 0.5

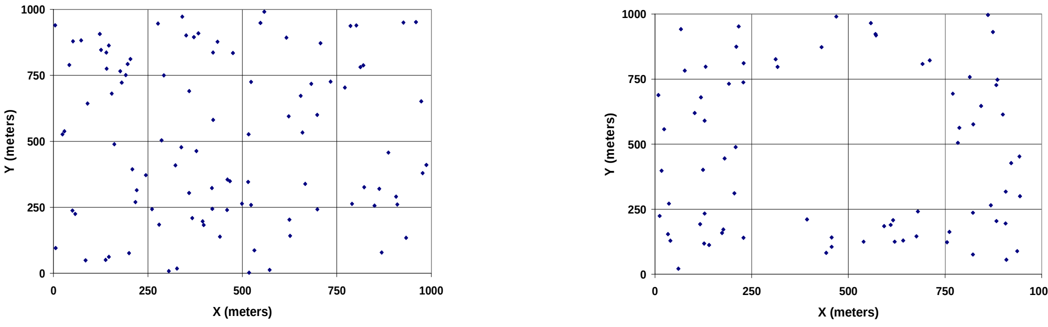
(b) Packet Rate = 1.0

Figure 13. Location of hotspots for Random Scenarios

Hotspot nodes are marked with as big darker circles. A plot of the positions of hotspot nodes indicates that for random scenarios, these nodes are found to be generally in vicinity of the center of the simulation area. It can be seen that hotspots occur generally within a square of side 500 meters centered at the center of simulation area i.e. (500,500). The percentage of packets dropped at hotspot nodes is 44.59% and 45.35% for (a) and (b) respectively.

For uniform layout due to random distribution of nodes, the connection pairs are generally across the network resulting in higher average hop count. Thus most of the time data packets from the original sender reach the destination with the aid of intermediate forwarding nodes. In random scenarios nodes in the vicinity of the center of simulation area, are expected to forward most of the data packets. This is due to the fact that most of the routing protocols intrinsically use the shortest path algorithm to find the routes, which is usually the center in random scenarios. Thus the intermediate nodes are unevenly burdened compared to the nodes along the edges of simulation area.

This fact was further confirmed with a simple simulation setup wherein the center of the simulation doesn’t exist or in other words there are no nodes present in the center. Thus the routing protocol is forced to choose routes away from the center. For this purpose the 4 sub-areas having one of their edge points as the center were kept empty and the nodes were placed randomly in the remaining areas. This is shown in Figure 14. To keep the node density proportional due to loss of 4 sub-areas (500x500 m^2) area in this case), the number of nodes were reduced to 75.



(a) Random Scenario (b) Random Scenario w/o center

Figure 14. Location of Nodes within Sub-Areas

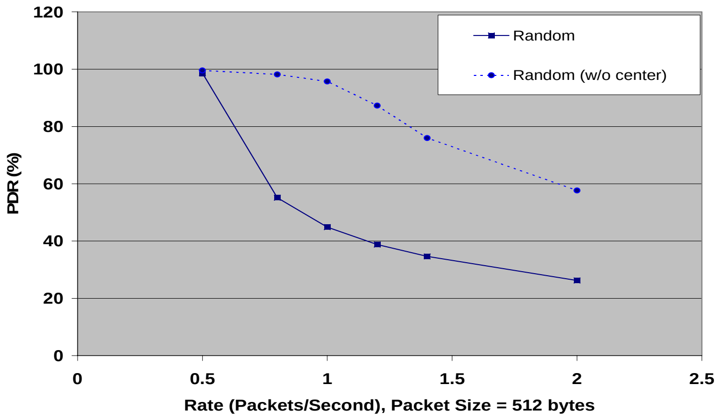


Figure 15. PDR Comparison for Random and Random w/o center

The PDR comparison for the two random scenarios is shown in Figure 15. One can see that as the bottleneck in random scenarios (nodes in vicinity of center) are removed, the performance improves drastically. Also the random layout without the center has a much lower delay and throughput. This is because due to the symmetry of the simulation area and assuming that nodes and traffic connection pairs are uniformly distributed over the simulation area, the traffic tends to be evenly distributed among all nodes.

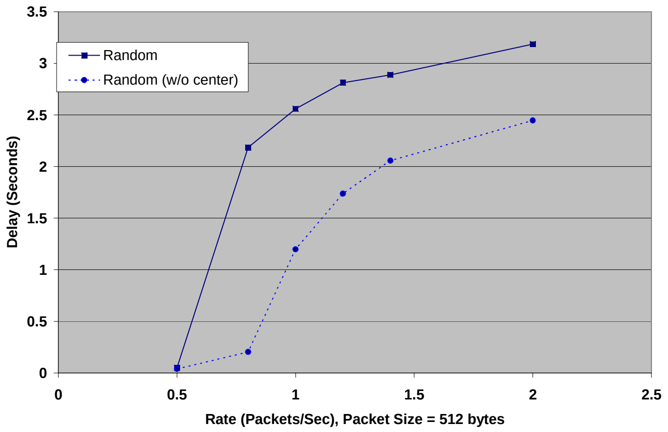


Figure 16. Delay Comparison for Random and Random w/o center

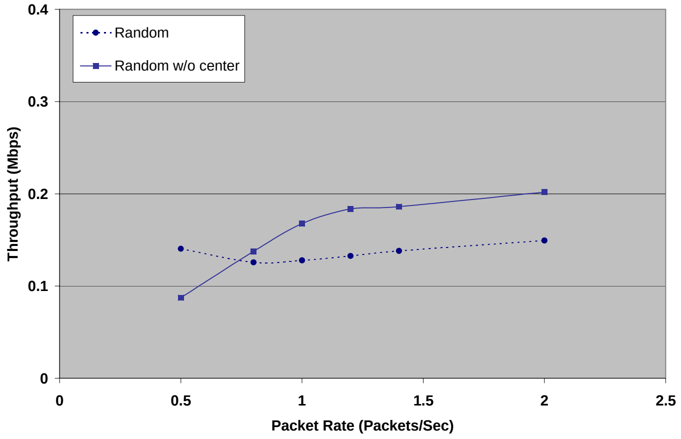
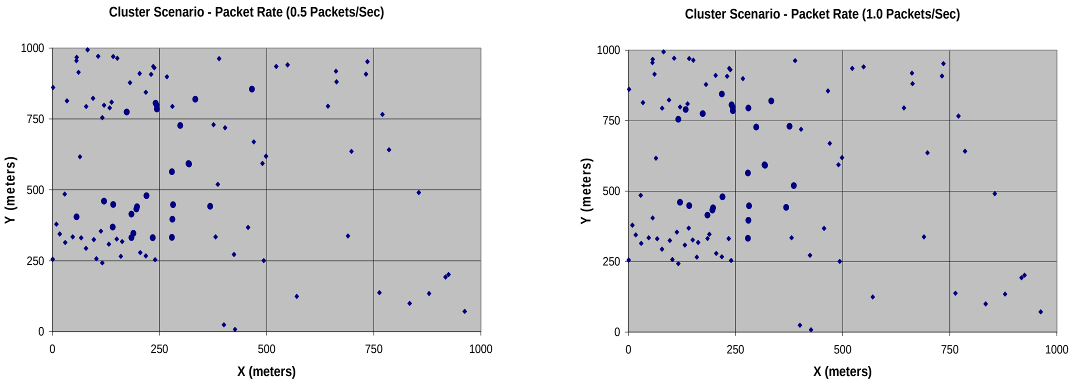


Figure 17. Throughput Comparison for Random and Random w/o center

These arguments support the views/statement in [21] that the random layout with random traffic distribution is a pessimistic scenario.

Hotspot Locations in Clustered Scenarios

For clustered layout defining the region of hotspot nodes becomes a difficult task since location of hotspot nodes depends on the location of the clustered sub-areas. However, for most cases it has been found that the hotspots are usually formed at nodes which act as a bridge for the inter cluster traffic. Such hotspots are distributed according to location of clustered regions and also the traffic pattern. Location of hotspots for a clustered scenario for 0.5 and 1.0 packet rates (intra cluster traffic – 30-40%) is shown below in Figure 18.



(a)Packet Rate = 0.5 (b)Packet Rate = 1.0

Figure 18. Location of hotspots for Clustered Scenarios

Hotspot nodes are marked as big darker circles. Percentage of packets dropped at hotspot nodes is 43.71% and 44.33% for (a) and (b) respectively.

The results presented in this thesis support the argument that the performance degrades due to existence of hotspots [29], defined as the nodes, which experience excessive contention, congestion and resource depletion. Hotspot nodes become the bottleneck in the network performance. Hotspots are generally created when traffic *converges* to a node or *small* cluster of nodes [29].

CHAPTER V

CONCLUSIONS

Mobility models play an important role in performance analysis studies of ad hoc networks. In this light, a realistic mobility model called as the Clustered Mobility (CM) model has been proposed this thesis. Non-homogenous node distribution has been characterized and modeled using an algorithm based on preferential attachment mechanism used modeling scale-free complex networks. The CM model has been implemented in Ns-2 and verified to produce a heavy-tail distribution for node densities in static scenarios. A definition of cluster, inter cluster traffic and intra cluster traffic has also been provided.

A performance evaluation of static wireless multi hop networks in presence of node clustering was done with aid of the CM model. It has been found that the performance of clustered scenarios varies significantly depending on the traffic pattern. In a clustered scenario the traffic pattern can be assumed to be made of intra cluster and inter cluster traffic. In general, larger percentage of intra cluster traffic yields higher throughput and packet delivery ratio with lower delays. This is due to the fact that with larger intra cluster traffic average hop count decreases which in turn reduces the forwarding load and routing overheads. Also, for a given percentage of intra cluster traffic a higher degree of node clustering degrades the performance. This is due to higher contention for the available bandwidth and lowered spatial reuse.

Increasing traffic intensity resulted in performance drop for both random and clustered layouts. A detailed study of the number of packet drops led to the conclusion that the performance degraded mainly due to the presence of hotpots. Hotspots are formed when traffic converges to a node or a small cluster of nodes. This result in an uneven loading of these nodes compared to other nodes in the simulation. These nodes suffer from high congestion and contention and thus drop a significantly more data packets compared to other nodes.

A position plot of the hotspot nodes for random layout revealed that such hotspot nodes generally lie in the vicinity of the center of simulation area, since for random scenarios center of simulation area yields in the shortest path metric for routing protocols. A similar plot for multiple clustered scenarios led to the conclusion that, in such scenarios hotspots are usually formed at nodes which act as a bridge for the inter cluster traffic. Such hotspots are distributed according to location of clustered regions and also the traffic pattern.

While the present model is able to produce realistic node distributions, it remains an important future work to incorporate realistic node movement and identifying probable hotspot areas with node mobility. Node speed and movement path are dependent on a range of interacting factors. Incorporating such factors into the model requires a diverse and in depth knowledge of various subjects.

BIBLIOGRAPHY

1. Albert-László Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, 288(50-59), 2003.

2. Amit Jardosh, Elizabeth M. Belding-Royer, Kevin C. Almeroth, and SubhashSuri. Towards realistic mobility models for mobile ad hoc networks. In *Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 217-229, 2003.

3. Amotz Bar-Noy, Ilan Kessler, and Moshe Sidi. Mobile users: to update or not to update? *Wirel. Netw.*, 1(2):175-185, 1995.

4. Ben Liang and Zygmunt J. Haas. Predictive distance-based mobility management for multidimensional PCS networks. *IEEE/ACM Trans. Netw.*, 11(5):718-732, 2003.

5. Biao Zhou, Kaixin Xu, and Mario Gerla. Group and Swarm Mobility Models forAd Hoc Network Scenarios Using Virtual Tracks. In *Proceedings of MilCom '04,Monterey, California, USA*, 2004

6. Chiang. *Wireless Network Multicasting*. PhD thesis, University of California, Los Angeles, 1998.

7. C. Perkins. Ad-hoc on-demand distance vector routing, *In MILCOM '97 panel on Ad Hoc Networks*, Nov. 1997.

8. C. Yu, K. G. Shin and B. Lee. Performance impact of node clustering in mobilead hoc networks. *International Conferenceon Wireless Networks and Emerging Technologies (WNET 2004)*, 2004. (submitted).

9. Christian Bettstetter, Hannes Hartenstein, and Xavier Pérez Costa. Stochastic properties of the random waypoint mobility model. *Wirel. Netw.*, 10(5):555-567, 2004.

10. David B Johnson and David A Maltz. Dynamic Source Routing in Ad Hoc Wireless networks. In Imielinski and Korth, editors, *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.

11. Elizabeth M. Royer, P. Michael Melliar-Smith, and Louise E. Moser. An Analysisof the Optimum Node Density for Ad hoc Mobile Networks. In *Proceedings of the IEEE International Conference on Communications, Helsinki, Finland,* June 2001.

12. F. Bai, N. Sadagopan, A. Helmy. The IMPORTANT Framework for Analyzing the Impact of Mobility on Performance of Routing for Ad Hoc Networks. *AdHoc Networks Journal - Elsevier Science, AdHoc Networks Journal - Elsevier Science*, 1, Issue 4:383-403, November 2003.

13. http://may.cs.ucla.edu/projects/maisie/models.html. Accessed on, Nov 22, 2004.

14. http://pcl.cs.ucla.edu/projects/glomosim/. Accessed on, Nov 22, 2004.

15. http://web.informatik.uni-bonn.de/IV/Mitarbeiter/dewaal/BonnMotion/. Accessed on, Nov 27, 2004.

16. http://www.isi.edu/nsnam/ns/. Accessed on, Nov 21, 2004.

17. http://www.webopedia.com/term/h/hotspot.html. Accessed on, Nov 29, 2004.

18. http://www.websters-online-dictionary.org/definition/ad+hoc. Accessed on, Nov 16, 2004.

19. Ingmar Glauche, Wolfram Krause, Rudolf Sollacher, Martin Greiner. Continuum percolation of wireless ad hoc communication networks. *Statistical Mechanics;Disordered Systems and Neural Networks, Journal-ref: Physica A 325 (2003) 577—600.*

20. J. Yoon, M. Liu, and B. Noble. Random Waypoint Considered Harmful, In*Proceedings of INFOCOM. IEEE,* 2003.

21. Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee, and Robert Morris. Capacity of ad hoc wireless networks. In *Mobile Computing and Networking*, pages 61-69, 2001.

22. Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu, and Jorjeta Jetcheva. A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing protocols. In *Mobile Computing and Networking*, pages 85-97, 1998.

23. Juan Carlos Cano, Pietro Manzoni, Miguel Sanchez. Evaluating the impact of group mobility on the performance of mobile ad hoc networks. *2004 IEEE International Conference on Communications (ICC 2004), Paris, France*, pages 20-24, June 2004. (to appear).

24. K. Wang and B. Li. Efficient and Guaranteed Service coverage in partitionable mobile ad-hoc networks, In *Proceedings of IEEE INFOCOM*, 2002.

25. K. Wang and B. Li. Group Mobility and Partition Prediction in Wireless Ad-Hoc Networks, In *Proceedings of IEEE International Conference on Communications(ICC), NYC, NY*, April 2002.

26. M.J. Feeley, N.C. Hutchinson, S. Ray. Realistic mobility for mobile ad hoc network simulation. *In Proceedings of the Third International Conference on Ad-Hoc, Mobile and Wireless Networks*, July 2004.

27. Miguel Sánchez and Pietro Manzoni. ANEJOS: a java based simulator for ad hoc networks. *Future Gener. Comput. Syst.*, 17(5):573-583, 2001.

28. P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory* , vol. IT-46, no. 2, pp. 388-404, March 2000.

29. S. Lee and A. Campbell. HMP: Hotspot Mitigation Protocol for Mobile Ad Hoc Networks, In *11th IEEE/IFIP International Workshop on Quality of Service, Monterey, Canada*, June 2003.

30. Samir Ranjan Das, Charles E. Perkins, and Elizabeth E. Royer. Performance comparison of two on-demand routing protocols for ad hoc networks. In *INFOCOM (1)*, pages 3-12, 2000.

31. T. Camp, J. Boleng, and V. Davies. A Survey of Mobility Models for Ad Hoc Network Research. *Wireless Communications Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5):483-502, 2002.

32. V. KAWADIA and P. KUMAR. Power control and clustering in ad hoc networks, In *Proc. IEEE Infocom (San Francisco, California*, Apr. 2003.

33. W. Navidi and T. Camp. Stationary distributions for the random waypoint mobility model. *IEEE Transactions on Mobile*, 3(1):99-108, January-March 2004.

34. X. Hong, M. Gerla, G. Pei, and C. Chiang. A Group Mobility Model for Ad Hoc Wireless Networks. In *Proceedings of ACM/IEEE MSWiM'99, Seattle, WA,* pp.53-60., Aug. 1999.

35. Yu, C. Shin, K.G. Lee, B. Power-stepped protocol: enhancing spatial utilization in a clustered mobile ad hoc network. *Selected Areas in Communications, IEEE Journal*, 22, Issue: 7(0733-8716):1322- 1334, Sept. 2004.

36. Z. Haas. A new routing protocol for the reconfigurable wireless networks. In *Proc. of the IEEE Int. Conf. on Universal Personal Communications.*, Oct 1997