Understanding the performance limitations of a simulated large scale wireless ad-hoc network

Penina Orenstein
Seton Hall University
Department of Computing and Decision Sciences
South Orange, New Jersey, USA
orenstpe@shu.edu

Zory Marantz
The New York City College of Technology
Department of Electrical Engineering
Brooklyn, New York, USA
zory@ieee.org

Abstract

We use a cellular automaton analogy to simulate a large scale wireless ad-hoc network. This approach combines mobility in its simplest form with some fundamental attributes of radio propagation and enables us to examine the communicative properties of the network which would not otherwise be accessible. The analysis shows that there is an optimal network density for which the throughput of the network is maximized. We examine this finding under a range of processing gain values and confirm (a) that both the maximum total network throughput and the network's sustainability increases proportionately with the processing gain and (b) that single-hop communication is always preferable to extended-hop communication. Furthermore we consider the performance of the system when communication is not limited to a single-hop. We show how processing gain can be used adaptively in order to control the transmission range and hence guarantee end-to-end connectivity in the network.

1. Introduction

In a preliminary study a simple cellular automaton model [1,2] was proposed to study the key aspects of mobility (both free and restricted), on the transfer rate of information. As part of the study, the authors' developed a generic wireless network whose topology and radio properties emerged as a result of an underlying cellular automaton model, [3]. Some simplifying assumptions were made, yet the model revealed qualitative features of wireless communication which would not have otherwise been accessible. Specifically, we showed how mobility enhanced communication across a range of network densities and also demonstrated the benefits of short range communication in terms of improved network throughput for a fixed processing gain G. This paper builds on these findings, in particular, by examining the impact of variable processing gains. Consequently, we first summarize the mechanism of the underlying model and point out the key parameters. We then use the model to demonstrate the contribution of the processing gain parameter.

2. An Integrated Model

2.1. Model Overview

The model simulates the movement of nodes in a grid network. The MATLAB software suite was used for the simulation environment. Effectively, we create a large matrix (typically set at 20x20, but this can be adjusted) which keeps track of the location of each node in the grid and referenced using the (i,j) position in the grid. At each time-slice, a node moves to its next location where the probability of movement is determined by a system parameter, p_{stay} – the probability that a node stays in its current location. This probability represents the rate at which terminals move around and simulates a random walk through the network, [2,4]. The simulation begins with an initial setup phase which populates the grid with the nodes. This is followed by the simulation phase during which the nodes move around the grid and simulation statistics are recorded (See Table 1).

A central theme in our model is the concept of a *neighbourhood*. For any given node S at a given moment, there are three distinct types of neighbourhood:

a) The *mobility neighbourhood*: the set of possible locations that the vehicle can move to during any one time slice (N,S,E,W).
b) The *communication neighbourhood*: the set of nodes that are in its range, and
c) The *interference neighbourhood*: the set of nodes that contribute towards interference with radio reception for that node.

As one might expect, in our model, the communication neighbourhood is larger than the mobility neighbourhood: the latter includes only the four cells immediately bordering the cell occupied by node S (without off-diagonal positions), whereas the communication neighbourhood additionally includes the four diagonally neighbouring cells NW, NE, SW, SE. We shall refer to this as *single-hop* communication. In principle, with a more powerful transmitter, node S could transmit directly to vehicles further away, in non-neighbouring cells, a point that will be explored later. We shall refer to this as multi-hop communication. Note that in this paper, the number of 'hops' refers to spatial separation; other authors use

it in a different sense, to mean the number of stages in relaying a message between node S and node D.
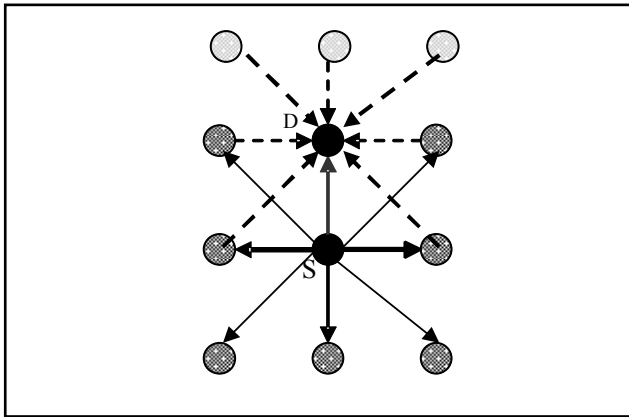
Figure 1. Calculating SNR requirements. Solid black node (S) can communicate with any node within its neighbourhood (shown via solid lines). Dashed lines emanating from nodes arriving at D, contribute towards the interference at node D.

In a study of wireless transmission in an ad hoc network, it is necessary to make some assumptions about the information 'load': the number of messages to be transmitted or relayed by each node per unit time. Here, we assume that the rate is fixed independently of other factors such as the density of nodes on the network. Each message is broken down into packets, and then each packet is routed independently via the route with the strongest signal strength. The total throughput achieved within the network as a whole is defined as the number of packets successfully transferred per second, and this is one of the two parameters that we use as overall measures of performance.

2.2. Model Parameters

The operation of the simulation model is controlled by two sets of parameters, the first deals with the topological properties of the network, and the second which relates to wireless communications. A list of these parameters is provided for reference in Table 1, together with two output parameters that we use as measures of the overall performance of the wireless system. A more detailed description of these parameters is given below together with the derivations for the formulae used to derive some of the results described in this paper.

We have assumed that $P_i(t) = P_{max} = 1$ (in watts) for all i, so that each node emits signals at a fixed, maximum power level. This corresponds to the worst case for interfering communications. When a radio signal is emitted, like almost all forms of radiation its strength declines with distance away from the source. Here, we assume that the strength of the signal received at distance d is given by

$$P_{rec} = \frac{cP_{max}}{d^{\alpha}} \qquad (1)$$

where α is the path loss exponent. The multiplicative constant c is related to the real world requirement to normalize some parametric values. Initially it was set to 1.

Our model only considers large-scale path-loss characteristics, i.e., the power of the signal declines with distance between emitter and receiver according to a specified functional relationship independently of local topography. The model does not take into account more subtle effects such as reflection and refraction leading to multiple signals (multipath fading), or shadowing effects caused by obstructions such as hills or buildings.

Table 1. Key parameters in the model

	Symbol	Definition
	p_{stay}	Probability that a vehicle remains in its current location during any time slice
GRID NETWORK PARAMETERS	μ	Mean total demand in nodes per time slice (Poisson arrivals, distributed around the grid perimeter)
	N	Size of grid (in this study, fixed at 20, leading to a grid of 20 x 20 = 400 cells)
	M	Mobility neighbourhood (since nodes can only move between neighbouring cells, M is set to 1 in this study)
	$N_T = 1,2,\ldots$	Transmission neighbourhood (includes diagonally neighbouring positions). The value is larger than 1 for multi-hop communication.
	$N_R = 1,2,\ldots$	Interference neighbourhood (includes diagonally neighbouring positions). The value is larger than 1 for multi-hop communication.
WIRELESS NETWORK PARAMETERS	$P_i(t)$	Power of node i at time t, initially set to $P_{max} = 1$ (watts). This corresponds to the worst case for interfering communications.
	α	Path loss exponent
	d	Geographical distance between nodes (units of the grid)
	G	Processing gain ($G>1$) See definition below
	$\sigma^2 = 5 \times 10^{-15}$	Additive White Gaussian Noise (AWGN) in watts
	R	Normalized data rate equal to 1 packets/sec
SIMULATION PERFORMANCE PARAMETERS	Φ	Wireless node density = no of occupied cells/n^2
	TN	Total network throughput = $R \times$ (total number of allowable connections)

The power of the signal emitted by node i and received by node j is $P_i(t)g(d_{ij}(t))$, where $d_{ij}(t)$ is the geographical distance between nodes i and j on the grid at time t, and $g(\tilde{d})$ is the channel gain function in the wireless medium, with $\tilde{d} = |d_{ij}(t)|$.

At time t, node i transmits data to node j provided the signal-to-interference ratio SNR is larger than a threshold β. The SNR is found via

$$SNR = \frac{GP_{rec,\,centre}}{(P_{rec,\,other} + AWGN)}$$
(2)

where G is the processing gain, defined as the ratio between system bandwidth (W Hz) and R the channel bandwidth, and

$$P_{rec,\,center} = P_i(t)\big(g(d_{ij}(t)\big)$$
(3)

$$P_{rec,\,other} = \sum_{\substack{k \neq i,\,j \\ k \in N}} P_k(t)\big(g(d_{kj}(t)\big)$$
(4)

this last quantity being the interference contribution from nodes within the neighbourhood of the receiving node j.

As shown in Figure 1, we assume that node S successfully transmitted data to node D if the signal received by D is greater than a randomly generated threshold, β, (0 ≤ β ≤ 1), in which case a fixed amount of information (R packets/s) flows between the node pair. If there is more than one possible candidate connection, then the stronger one is used.

We use a mathematical function f(SNR) to encapsulate the frame success function - the probability that a node's data packet is received successfully, without errors at the decoder. The dependent variable is the received signal-to-interference ratio SNR. The transmission takes place providing f(SNR) ≥ β. The specific form of the function f(SNR) depends on the details of the transmission system, such as, modem configuration, channel coding, antenna configuration, and radio propagation conditions. Our analysis applies to a wide class of practical frame success functions, each characterized by an S-shape form [5].

At the end of each simulation cycle, the model records (a) the total number of established connections, (b) the total normalized network throughput TN, which is the total number of connections multiplied by the data rate of R packets/s, and (c) the wireless node density, referred to as Φ: the number of occupied sites divided by the maximum number of possible sites in the grid.

2.3.   Relation to other work

There is a vast amount of theoretic research which deals with mobility and throughput capacity of wireless ad-hoc networks, see [6-9] and the references therein. Two fundamental papers in this area include, the work by [10] and by [11]. In [10], the authors propose a model to study the capacity of fixed ad-hoc networks, where nodes are randomly located but are immobile. The main result shows that as the nodes per unit area n increases, the throughput per source-destination pair decreases approximately like 1/√n. The fundamental performance limitation comes from the fact that long range direct communication between many user pairs is infeasible due to excessive interference caused by nodes in the vicinity and so, most communication has to occur between nearest neighbors.

In [11], mobility is introduced to overcome this limitation, so that two nodes communicate only when the source and destination nodes are close together. This resembles the Infostation architecture, [12], where users connect to the infostation only when they are close by. The authors demonstrate that the average long-term throughput per source-destination (S-D) pair can be kept constant even as the number of nodes per unit area increases. This improvement stems from the time variation of the users' channels due to mobility. The authors define an optimal network density which maximizes the network throughput and demonstrate how throughput increases with density but only up to a point, whereupon the network becomes overpopulated and throughput begins to tail off. The results pertain to a narrowband system (where the processing gain is 1). Also, the concept of "restricted" mobility is not dealt with explicitly.

This paper uses simulation (a) to confirm the key findings in [10-11] and (b) to extend these findings to the spread-spectrum case (i.e. where the processing gain is larger than one). We then build on these results and show how the processing gain G can be used to design effective coverage areas which maximize the total network throughput for a range of network densities.

3.   Results

3.1.   Effect of wireless node density

The results in Figure 2 show that as the wireless node density increases, so does the normalized network throughput, but only up to a point. The maximum throughput, which we denote by TN*, occurs at a particular value of the node density, which is denoted by Φ_{opt}. In the case of single-hop communications, where each device may communicate with one of eight nodes in its immediate neighbourhood, this maximal throughput occurs when $\Phi_{opt} \approx 0.30$. Increasing the number of nodes has the effect – at first - of promoting communication, but in congested situations the wireless network effectively shuts down because of increased overcrowding (and therefore interference) among the nodes.

Multi-hop communication only makes matters worse. If, for example, a device can 'talk' to nodes located two hops away, the level of interference increases and the total network throughput is significantly reduced. The optimum node density Φ_{opt} is also reduced, in this case to a value of roughly 0.13. The optimum node density is even less for three-hop communication.
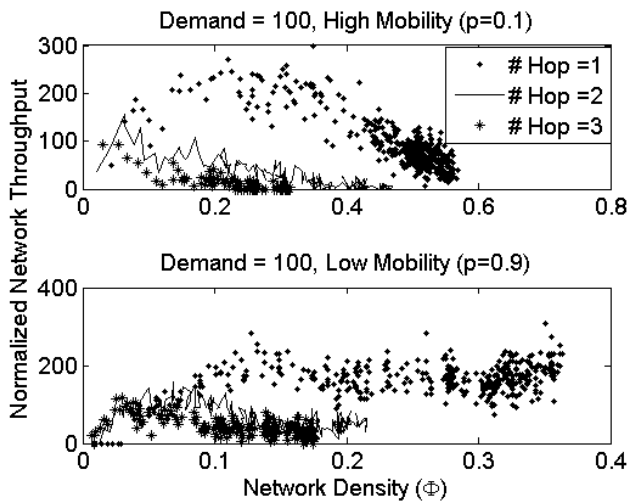
3.2. Impact of Mobility

Communication is also affected by mobility. Again, referring to Figure 2, with low mobility, no communication is

possible once the network reaches a density of $\Phi \approx 0.35$. By contrast, with high mobility, some throughput (albeit suboptimal) is possible up to network densities of $\Phi \approx 0.55$. Table 2 summarizes the values of Φ_{opt} and TN^* for selected mobility scenarios. Even with quite high node densities, communication can still take place provided the nodes are mobile.

Table 2: Impact of communication/interference range on critical network density $_{opt}$ and corresponding maximal throughput TN*

Mobility Level	No. of hops	Φ_{opt}	TN* (packets/s)	Density Φ for which network fails irretrievably
High	1	0.30	250	0.55
Low	1	0.13	200	0.35
High	2	0.10	150	0.48
Low	2	0.08	150	0.22
High	3	0.07	100	0.31
Low	3	0.07	100	0.18

Figure 2. Effect of Mobility on Communication with fixed processing gain G=10. Mobility extends network sustainability and single-hop communication maximized network throughput with $_{opt}$=0.30 for high mobility and $_{opt}$=0.13 for low mobility.

At first sight this may seem surprising. The explanation lies in the fact that congestion implies wireless interference. In our treatment, low mobility is equated specifically with a high value of p_{stay}, in other words, a high probability of a node being forced to remain in position by external factors (such as blocking) rather than moving to an adjacent site, during any given time slice. If a large proportion of the nodes are held up in this way, these nodes in turn will obstruct other nodes and generate local queues: high concentrations or pockets of nodes that are blocking each others' way. This in turn implies a high proportion of nodes whose messages are subject to interference

through local overcrowding, and hence a reduced wireless throughput.

3.3. Variable Processing Gain

In general, increasing the processing gain (a scarce resource) improves the total network throughput which is maximized at a particular network density. Figures 3a-3c shows that when the processing gain is increased, (a) the total network throughput increases proportionately, and (b) the feasible communication range is extended. We summarize the critical values of Φ_{opt} and TN^* for each processing gain level and communication range in Table 3. The data suggests that for single-hop communication the optimal network throughput increases proportionately with higher processing gains and that communication is possible as long as occupation density is less than 70%. The graphs in Figures (3b-3c) indicate similar trends but they highlight the point that with multi-hop communication, both the maximal throughput and network sustainability levels are considerably reduced.

Table 3: Effect of increasing processing gains on communication range and total network throughput level

	G	Φ_{opt}	TN* (pkts/sec)	Density Φ for which network fails
Single-hop	20	0.35	1000	0.70
	30	0.48	3100	0.70
	40	0.60	6500	0.70
Two-hop	20	0.10	500	0.48
	30	0.30	1750	0.58
	40	0.40	3900	0.60
Three-hop	20	0.05	100	0.47
	30	0.25	1250	0.50
	40	0.35	2800	0.47

By comparing the first and last rows of Table 3, one can observe how two optimal densities (0.35) are achieved at different levels of processing gain, G=20 (single-hop) and G=40 (triple-hop). Correspondingly, their critical density is 0.7 and 0.47 respectively. This information can be used when prioritizing certain data streams. In order to transmit important data which needs to reach the destination quickly, the network will have to (a) operate at the higher processing gain (b) transmit further away (albeit with the possibility of collapse sooner) yet a higher throughput will be achieved. Data streams with lower priority should operate at a lower processing gain, but will be connected for longer periods and achieve a lower throughput.

These results suggest that it may be possible to use the processing gain adaptively depending on a network administrator's performance goals. For example, suppose a network administrator requires TN*=3000 packets/s, one can allow for communication across two-hops providing G=40 or, one can restrict communication to single-hop mode by using G=30. The latter design yields Φ_{opt}=0.51 vs. Φ_{opt}=0.35 for

two-hop communication. The preferred mode will depend on network performance goals and current traffic conditions. That is, if nodes are sparsely located, single-hop communication might not be possible, and extended-hop communication might be the only way of maintaining network connectivity.
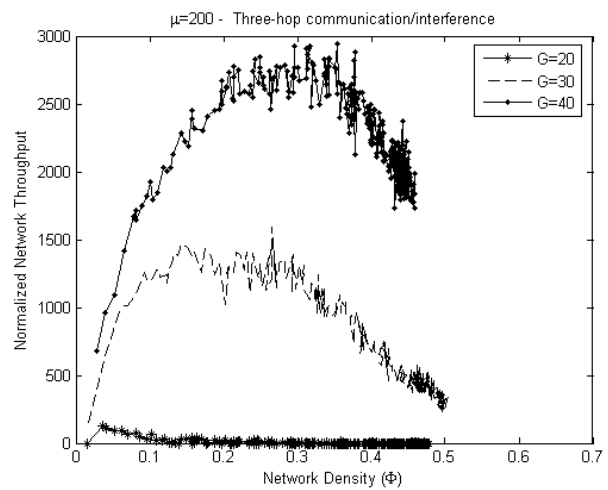
The graphs in Figure 3a-3c can be used to determine the best mode of operation. The results show that with limited processing gain, single-hop communication is always superior to multi-hop communication. However, depending on the occupation density, one might need to adjust the processing gain in order to provide coverage over wider areas of the network.
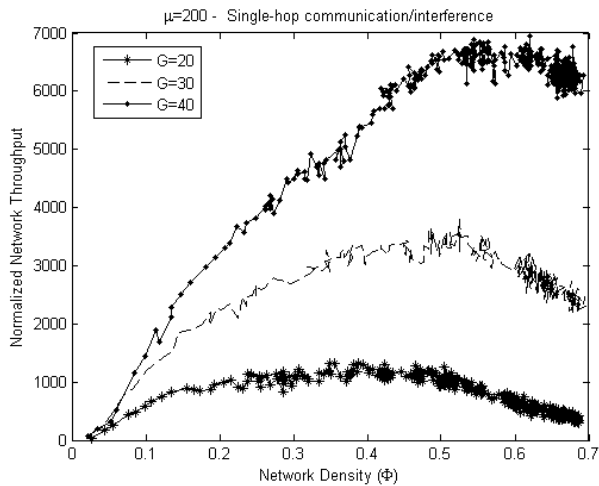
(a)
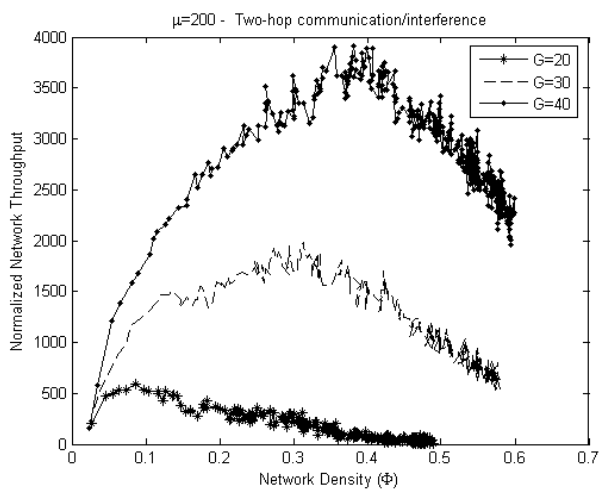
(b)

(c)

Fig. 3(a-c): Effect of variable processing gain on total network throughput and optimal network density for different ranges of communication. Graphs can be used to select suitable operating points depending on network conditions.

The results of this simulation point to a dynamic approach which should be applied in order to maximize the throughput in the network. The elasticity of the communication range needs to be linked to the prevailing network conditions. Thus, given the largest available processing gain, G, one can allow for long-range communication when the density of nodes is low, but as this density increases, the range of communication needs to be reduced in order to maintain service quality goals (maximal throughput and network sustainability).

4.   Conclusions and Further work

In this paper, we have proposed a cellular automaton based simulation model to study some of the emergent properties of a generalized mobile ad-hoc network. The analysis points to an interdependent relationship between the spatial distribution of nodes, communication and interference. The main results can be summarized as follows:

a) Given a number of nodes in a particular region, they will be more evenly spaced if they are moving freely. By contrast, under congested conditions, they will be spaced unevenly, with some locked in tight little knots such that interference inhibits communication. Mobility is the cause for improved communication because of the resulting spatial distribution.

b) At a particular density of nodes, the total network throughput is optimal but falls beyond this critical density.

c) The processing gain parameter can be used to control the communication range as well as to increase the level of throughput. With limited processing gain, single-hop communication achieves highest throughput levels and our results suggest that the processing gain should be elevated in order to extend

the coverage region across wider areas of the network or to boost the total throughput. The processing gain can also be used adaptively for prioritized traffic streams.

Our simulations suggest a flexible topology of wireless ad-hoc networks which can be controlled via adaptive processing gain levels. We demonstrated that in order to maximize network throughput and increase network survivability, administrators should tailor the communication range according to the prevailing network conditions. Longer range communication is appropriate only when the network density is low, but as the network density increases this range should be reduced proportionately. One way of achieving adjustable communication ranges is by increasing the processing gain according to network density conditions.

This paper has dealt only with conceptual issues arising from communication in a generalized form of mobile ad-hoc network, but the results can be viewed in the context of a specific form of ad-hoc network, namely a VANET (vehicular ad-hoc network). The model described here has been used to investigate the relationship between communication, contention/interference and mobility in a generalized context, but these features are all of primary concern in a VANET where congestion among nodes and the impact on communication is a cause for concern. In addition, both the quality of communication across the VANET as well as the end-to-end connectivity are both important issues to be explored. Our observations regarding single-hop vs extended-hop communication will bear an impact when designing such networks. For example, in a sparse VANET, it may be necessary to introduce extended-hop communication in order to maintain end-to-end connectivity. Even though the overall network throughput will be less than that would be achieved through single-hop communication, nevertheless, this would be required in order to satisfy the end-to-end connectivity requirement.

Among the applications for a VANET so far suggested are the propagation of safety warnings such as icy road conditions, crime prevention, surveillance aimed at public security, together with less urgent passenger services, and even congestion management. To determine whether and how such a system would function, it is necessary to model two distinct kinds of network simultaneously – the road system and the wireless network. The challenge is significant, not least because of the many factors involved. The model in this paper forms the baseline for our future work, but we will need to modify the somewhat crude mobility model to reflect the vehicular urban environment.

We are aware that our results from this study are subject to caveats arising from the simplistic nature of our assumptions, both in terms of node mobility and in terms of the likely pattern of demand for communication in real systems. On urban road networks, vehicles tend to move around in 'platoons', that continually expand and contract as they move through obstructions such as traffic signals. For some of the time, vehicles are close together, but there are frequently long gaps that signals cannot bridge. Moreover the demand for signal processing is likely to vary with the number of vehicles on the network, in ways that depend on the application envisaged. Nevertheless, the qualitative features we have observed point towards aspects of VANET behaviour that merit investigation with a revised, more realistic model. This is the subject of our future research.

5. References

[1] P. Orenstein and Z. Marantz, "Improving communication in wireless networks via mobility", *Proc*. INFORMS 2007, Puerto Rico, July 8th -11th, 2007.

[2] P. Orenstein and Z. Marantz, "Application of cellular automata to modeling mobility and radio wireless communication in wireless networks", *Proc.* IEEE SARNOFF Symposium, April 2005.

[3] O. Biham and A.A Middleton "Self-organization and a dynamical transition in traffic flow models", *Physical Review* Vol. 46A, No. 10, pp 6124-6127, 1992.

[4] K.H. Chiang and N. Shenoy, "A 2-D Random-Walk Mobility Model for Location-Management Studies in Wireless Networks", *IEEE Transactions on Veh. Technology*, Vol. 53, No. 2, pp 413-424, March 2004.

[5] V. Rodruigez, Robust Modeling and analysis for wireless data resource management", IEEE WCNC, Vol 2, pp 717-722, March 2003.

[6] T. Camp, J. Boleng, V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research ", WCMC Special Issue on Mobile ad-hoc networking: research, trends and applications, 2002.

[7] J. Zander and S.L. Kim, Radio Resource Management for Wireless Networks, Artech House, Norwood, USA, 2001.

[8] T. Bonald, S.C. Borst, and A. Proutiere, "How Mobility Impacts the Flow-Level Performance," Proc. IEEE INFOCOM 2004,Hong Kong, China, pp. 1872-1881, 2004.

[9] C.T. Lea and M Zhang, "On the Mobility/Capacity Conversion in Wireless Networks," WCNC'03, Vol. 3, pp.1433-1438, 2003.

[10] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, Vol. 46 No. 2, pp 388-404, March 2000.

[11] M. Grossglauer and D. Tse , "Mobility increases the capacity of wireless ad-hoc networks", *Proc. IEEE INFOCOM*, Vol 3, pp 1360-1369, 2001.

[12] W.H. Yuen, R.D. Yates and S. Mau, "Exploiting Data Diversity and Multi-user diversity in noncooperative mobile infostation networks", *Proc. IEEE, INFOCOM 2003*.