Using Statistical Design of Experiments for Analyzing Mobile Ad Hoc Networks*

Michael W. Totaro
The Center for Advanced Computer Studies
University of Louisiana at Lafayette
Lafayette, LA 70504-4330
miket@cacs.louisiana.edu

Dmitri D. Perkins
The Center for Advanced Computer Studies
University of Louisiana at Lafayette
Lafayette, LA 70504-4330
perkins@cacs.louisiana.edu

ABSTRACT

The performance of mobile ad hoc networks can be influenced by numerous factors, including protocol design at every layer; parameter settings such as retransmission limits and timers; system factors such as network size and traffic load; as well as environmental factors such as channel fading. In this work, we are concerned with understanding the functional relationship between these influential factors and performance of mobile ad hoc networking systems. We show how a systematic statistical design of experiments (DOE) strategy can be used to analyze network system and protocol performance, leading to more objective conclusions valid over a wide range of network conditions and environments. Using a DOE strategy and a 2^k factorial design, we quantify the main and interactive effects of five factors (i.e., network density, node mobility, traffic load, network size, and medium access control scheme) on two response metrics (i.e., packet delivery ratio and end-to-end delay). Using these effects measures, we then develop two first-order linear regression models that define the functional relationship between the influential factors and two performance metrics.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: modeling techniques, performance attributes; G.3 [**Probability and Statistics**]: Experimental design, Correlation and regression analysis; I.6.6 [**Simulation and Modeling**]: Simulation output analysis

*This work was supported in part by the U.S. Department of Energy (DOE) under Award Number DE-FG02-04ER46136, by the State of Louisiana Louisiana Board of Regents under Contract Numbers DOE/LEQSF(2004-07)-ULL and LEQSF(2003-06)-RD-A-35, and by the National Science Foundation under NSF Career Grant No. 0448055.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MSWiM'05, October 10-13, 2005, Montreal, Quebec, Canada.
Copyright 2005 ACM 1-59593-188-0/05/0010 ...$5.00.

General Terms

Performance, Experimentation

Keywords

mobile ad hoc networks, statistical design of experiments, performance evaluation

1. INTRODUCTION

A Mobile ad hoc networks (MANET) is self-organizing, rapidly deployable, and do not depend on a predefined infrastructure such as base stations or access points [7]. Instead, ad hoc networks are comprised of wireless nodes (mobile or stationary) that must cooperate in order to dynamically establish communications using fully distributed protocols such as medium access control and adaptive wireless multihop routing. All of which must be done with limited network management and administration mechanisms. To address the inherent challenges of the ad hoc networking paradigm, researchers and engineers have been aggressively pursuing solutions at every layer of the protocol stack, ranging from routing at the network layer [5] to adaptive and distributed medium access control layer [11] to fair and robust protocols for a reliable transport layer [8].

As protocol design and development continues to mature and as the research community begins to consider standardization issues, interests must now focus more heavily on understanding performance and scalability tradeoffs and in order to make more objective conclusions when investigating and comparing alternative protocol and system architecture solutions. In general, several factors could potentially impact system and protocol performance. These factors can range from categorical factors such as protocols to quantitative factors such as network size, channel capacity, or transmission range. Moreover, preliminary work [4, 6] suggests that there exists significant cross-layer and parameter or factor interaction in the ad hoc networking environment, which has led to interest in a cross-layer or joint approach to protocol design as opposed to the traditional disjoint layered approach of the OSI and TCP/IP protocol architectures [13]. The large number of influential factor and cross-layer design stragegy leads to numerous questions: Which factors significantly influence a given performance metric? Is there any interaction among factors (e.g., protocols, retransmission limits, timers,)? Can we quantify the effects of each factor as well any potential interaction among factors? If so, which set of protocols and parameter settings provide optimal performance over a specific or wide range of operation scenarios? What is the appropriate choice of operating conditions to achieve desired performance?

In this work, we are concerned with gaining a better understanding of the functional relationship between these influential factors and system performance responses. In particular, our objectives are threefold: (1)identify the individual factors that signficantly impact a performance measure; (2) quantify the main and interactive effects of these factors; and (3) build mathematical models that can be used to charactize the relationship between the factors and various performance metrics. A common approach used to identify the impact of factors on one or more performance metrics is the traditional "one-factor-at-a-time," or OFAT, approach, where only *one* factor is varied, while holding all other factors constant. However, the OFAT approach does not allow us to consider the interaction of factors. That is, it does not consider that the influence of one factor may depend on the value of another factor. In effect, the OFAT approach is based on the assumption that the maximization value of one factor is independent of the level of the other; an assumption that usually is not true [3].

A useful analytical tool that can be used to quantify the effect of various factors on MANET performance is the statistical design of experiments approach. In this paper, we describe the use of a statistical DOE approach and show how this it can be useful in making more objective conclusions when investigating and comparing protocol performance. Furthermore, we develop empirical models for response variables crucial to the performance of ad hoc networks. The advantages of using a statistical DOE approach are significant. A proper DOE should allow the researcher to obtain the maximum information with the minimum of experiments [9]. Statistical DOE offers a way for the researcher to determine, *before* the runs are made, which specific configurations to simulate so that the desired information can be obtained with the least amount of simulating [12]. Additionally, an experimental design that has been properly executed and analyzed: (1) facilitates the identification of effects of various factors (variables) that might affect performance; and (2) helps to determine if a particular factor has a significant effect or if the observed difference is due to random variations that resulted from errors in measurement and uncontrolled parameters [9]. A complete tutorial on statistical DOE and empirical model building is outside the scope of this paper. For more in-depth study, please see texts by Box, Hunter, and Hunter [3]; Jain [9]; and Montgomery [14], and Kleijnen [10].

The remainder of this paper is organized as follows. In Section 2, we discuss both prior and current research and development work that focuses on the importance of empirically modeling performance in wireless ad hoc networks, as well as the requisite tools used by researchers that may help in developing these empirical models, useful primarily for protocol design and development. In Section 3, we describe in detail our experimental design, simulation setup, and data collection. In Section 4, we discuss the statistics generated, along with our analysis of these statistics. In Section 5, we describe the predictive models, derived from least-squares regression analysis, which can then be used to validate against expected results. Finally, we discuss our conclusions and future work in Sections 6.

2. RELATED WORK

Assessing the behavior of ad hoc networks is non-trivial. Nevertheless, empirical models that help to explain this behavior must be developed; thus, the primary objective of this paper. One of the most important tasks in developing empirical models is to identify key performance metrics and the factors that affect them. A set of performance metrics, useful for assessing MANET algorithms, is identified by Subbarao [16]; these include: average power expended, task completion time, end-to-end throughput, end-to-end delay, link utilization, and packet loss. In addition to these metrics, Subbarao describes *scenario metrics*, which include [16]: nodal movement/topology rate of change, number of network nodes, area size of network, density of nodes per unit area, offered load and traffic patterns, and number of unidirectional links. Compiling a list of factors that might influence system performance is inadequate. Thus, prior to accepting any set of factors as appropriate, we must determine whether or not they actually have any effect on the performance metrics of interest.

Vadde et al. [18]used statistical DOE to analyze the impact of factors and their interactions on MANET service delivery. Here the factors included: *QoS architecture*, *routing protocol*, *MAC protocol*, *offered load*, and *node mobility*. The performance metrics used to measure service delivered were: *real-time throughput*, *total throughput*, and *average delay*. Using statistical analysis of simulation data by way of analysis of variance (ANOVA) techniques, the researchers identified main effects and interaction of factors that explain the performance metrics. Vadde et al. [18] found that, for the average delay, the MAC protocol and its two-way interaction with the routing protocol are the most significant.

Analysis of variance (ANOVA) techniques were used by Barrett et al. (see [1] and [2]) to study empirically the effect of: (1) the interaction between the routing layer and the MAC layer in wireless radio networks; and (2) the interaction of the routing and MAC layer protocols using different mobility models. In the case of the former study, their analysis suggests that different combinations of routing and MAC protocols result in varying performance in different topology and traffic scenarios. Results of the latter study indicate that no single MAC/routing protocol combination dominated over all response variables, regardless of the mobility model used. These studies provide for a first step for the identification and analysis of main effects and factor interactions in wireless networks. Our work in this area is aimed at developing accurate mathematical models that characterize the relationship among the significant factors and multiple performance metrics. In this paper, we begin by illustrating how a 2^k factorial design can be used to determine the significant factors and then develop regression models based on the estimated main and joint effects for each factor.

3. STATISTICAL DOE

To yield objective conclusions, an experimental evaluation must comprise two key and interrelated components: (1) the experimental design, which refers to the process of planning the experiment so that data can be collected in a manner feasible for statistical analysis; and (2) the actual statistical analysis of the data [3, 14]. Our aim in this section is to provide a brief overview of statistical design of experiments

Table 1: Experimental Factors

Label	Factor	Level 1 (−)	Level 2 (+)
1	Avg. neighbors	7 (strongly-connected)	3 (weakly-connected)
2	Avg. node speed	5 m/sec (1-10 m/sec range)	30 m/sec (25-35 m/sec range)
3	Traffic load	10% of no. of nodes	20% of no. of nodes
4	No. of nodes	100	500
5	MAC layer	802.11b w/ RTS	802.11b w/out RTS

while introducing the specific experimental design and analysis techniques used in this study.

3.1 Terminology

Before we discuss our experimental strategy, it will be useful to define several standard DOE terms used in this paper [9]:

- *Factors:* The variables that affect the response variable. Factors may be classified as primary, secondary, or constant, depending on their use in an experiment design.
- *Levels:* The values that a factor can assume are called its levels.
- *Response Variable:* The measured performance of the protocol or system under study.
- *Design:* The experimental design specifies the number of experiments, the factor level combinations for each experiment, and the number of replications of each experiment.
- *Replication:* This refers to the process of repeating an experiment or set of experiments.
- *Main effects*[1]: Intuitively, the main effect of a factor refers to the average change in a response variable produced by a change in the level of the factor.
- *Interaction effects*[1]: Two factors *interact* if the performance response due to factor i at level m depends on the level of factor j. In other words, the relative change in the performance response due to varying factor i is dependent on the level of factor j.

3.2 Designing the Experiment

Step l: Defining the experimental objectives. Our underlying goal of this work is to demonstrate the effectiveness of a statistical DOE strategy when evaluating the performance of mobile ad hoc networking systems or protocols. To this end, the specific objectives of our experiments are to quantify the main and interactive effects of a subset of potentially influential factors on the performance of ad hoc networks. Using these effects, we then develop empirical models, which can be used to predict performance of the ad hoc system over the range of values examined in this work.

Step 2: Selecting the factors (and their levels). The next step in the experimental design process is selecting the potentially influential factors. In practice, numerous factors may impact the performance response of an ad hoc networking system. Since our overarching goal in this paper is to only illustrate the effectiveness of the statistical DOE strategy, in our prelimanary work, we have analyzed only a subset of five factors, while holding all other factors constant. Table 1 shows the factors studied in this current work.

We now provide justification for the factor levels (values) used in this study. *Average number of neighbors* is the average number of single-hop nodes within transmission range of any arbitrary node in the network. This can be considered a measure of network density and is expected to influence network connectivity, routing overhead, MAC contention, and source-destination path length and thereby influence the performance responses. For the average number of neighbors factor, we consider two levels: strongly connected (7 neighbors[2]) and weakly connected (3 neighbors). *Node mobility*, which is measured as the average node speed, will impact the frequency of topology changes. We also consider *traffic load*, which is measured as the percentage of nodes acting as source traffic generators. *Network size*[3], measured as the number of nodes in the system, will impact the path length and route discovery time, which could influence overall system performance. Finally, we consider the medium access control protocol as a primary factor. We investigate two levels: the IEEE 802.11 DCF with the optional RTS/CTS handshake and without RTS/CTS handshake. Research results show that the RTS/CTS handshake is useful in relatively static one-hop wireless networks. However, it is not clear what effect the RTS/CTS handshake will have in a multihop wireless environment with frequent topology changes where nodes move in and out of contention areas arbitrarily.

Step 3: Selecting the response variables. We consider two performance responses, each of which relates directly to the ability of the system to meet specific quality of service requirements. The *packet delivery ratio* is defined as the number of packets delivered to a destination divided by the number of packets actually transmitted. *End-to-end delay* is the application layer end-to-end delay, which includes all processing, queueing, and transmission delays at each node along the path.

Step 4: Selecting the appropriate design. We use a $2^k r$ factorial design. The $2^k r$ factorial design technique considers k **factors**, where each factor has two distinct levels (or values). For simplicity and computational purposes, it is often useful to code the factor levels as a + or − level, as shown in Table 2. The *design matrix* (see Table 2) shows all possible combinations of factor levels (called *design points*). Each design point corresponds to a simulation scenario, which is replicated $r = 5$ times, in our experiments. The response values for the performance metrics (i.e., *packet delivery ratio* and *end-to-end delay*) are also included in Table 2.

Step 5: Simulation and data collection. Our simulations were carried out using QualNet, a network modeling tool developed by Scalable Network Technologies. In order to obtain results that approximate an actual MANET, we ran each of the 32 simulations five times, after which we computed the average of the five runs for each design point. This results in a total of 160 simulation runs (32 de-

[1]A mathematical definition for our specific choice of experimental design is given in Section 4.

[2]It has been suggested by Kleinrock [17] and Royer [15] that throughput performance is optimal when the average number of neighbors is between six and eight neighbors.

[3]The terrain size was adjusted appropriately to maintain the required network density or average neighbors.

Table 2: Design matrix for 2^5 factorial design

		Factors					Performance Metrics	
		1	2	3	4	5		
Design Points	Factor Level	Avg. No. of Neighbors	Avg. Node Speed (m/s)	Traffic Load	Number of Nodes	MAC Layer	Packet Delivery Ratio	End-to-End Delay (secs)
	(−)	7	5	10% of Number of Nodes	100	802.11b w/ RTS		
	(+)	3	30	20% of Number of Nodes	500	802.11b w/out RTS		
1		(−) 7	(−) 5	(−) 10	(−) 100	(−) 802.11 b w/RTS	0.71568	0.86571101
2		(+) 3	(−) 5	(−) 10	(−) 100	(−) 802.11 b w/RTS	0.11592	1.27659734
3		(−) 7	(+) 30	(−) 10	(−) 100	(−) 802.11 b w/RTS	0.58568	0.9923993
4		(+) 3	(+) 30	(−) 10	(−) 100	(−) 802.11 b w/RTS	0.25776	2.13651797
5		(−) 7	(−) 5	(+) 20	(−) 100	(−) 802.11 b w/RTS	0.72484	0.76839629
6		(+) 3	(−) 5	(+) 20	(−) 100	(−) 802.11 b w/RTS	0.17076	1.41365995
7		(−) 7	(+) 30	(+) 20	(−) 100	(−) 802.11 b w/RTS	0.563	0.96332324
8		(+) 3	(+) 30	(+) 20	(−) 100	(−) 802.11 b w/RTS	0.24584	2.19733746
9		(−) 7	(−) 5	(−) 50	(+) 500	(−) 802.11 b w/RTS	0.395968	1.49277102
10		(+) 3	(−) 5	(−) 50	(+) 500	(−) 802.11 b w/RTS	0.092656	0.78984261
11		(−) 7	(+) 30	(−) 50	(+) 500	(−) 802.11 b w/RTS	0.271504	2.07584805
12		(+) 3	(+) 30	(−) 50	(+) 500	(−) 802.11 b w/RTS	0.08344	3.28247314
13		(−) 7	(−) 5	(+) 100	(+) 500	(−) 802.11 b w/RTS	0.330824	5.25921359
14		(+) 3	(−) 5	(+) 100	(+) 500	(−) 802.11 b w/RTS	0.099736	1.04019082
15		(−) 7	(+) 30	(+) 100	(+) 500	(−) 802.11 b w/RTS	0.16395	3.11871308
16		(+) 3	(+) 30	(+) 100	(+) 500	(−) 802.11 b w/RTS	0.07568	4.98781013
17		(−) 7	(−) 5	(−) 10	(−) 100	(+) 802.11 b w/out RTS	0.71568	0.865711017
18		(+) 3	(−) 5	(−) 10	(−) 100	(+) 802.11 b w/out RTS	0.11592	1.27659734
19		(−) 7	(+) 30	(−) 10	(−) 100	(+) 802.11 b w/out RTS	0.58568	0.9923993
20		(+) 3	(+) 30	(−) 10	(−) 100	(+) 802.11 b w/out RTS	0.25776	2.13651797
21		(−) 7	(−) 5	(+) 20	(−) 100	(+) 802.11 b w/out RTS	0.72484	0.76839629
22		(+) 3	(−) 5	(+) 20	(−) 100	(+) 802.11 b w/out RTS	0.17076	1.41365995
23		(−) 7	(+) 30	(+) 20	(−) 100	(+) 802.11 b w/out RTS	0.563	0.96332324
24		(+) 3	(+) 30	(+) 20	(−) 100	(+) 802.11 b w/out RTS	0.24584	2.19733746
25		(−) 7	(−) 5	(−) 50	(+) 500	(+) 802.11 b w/out RTS	0.395968	1.49277102
26		(+) 3	(−) 5	(−) 50	(+) 500	(+) 802.11 b w/out RTS	0.092656	0.78984261
27		(−) 7	(+) 30	(−) 50	(+) 500	(+) 802.11 b w/out RTS	0.271504	2.07584805
28		(+) 3	(+) 30	(−) 50	(+) 500	(+) 802.11 b w/out RTS	0.08344	3.28247314
29		(−) 7	(−) 5	(+) 100	(+) 500	(+) 802.11 b w/out RTS	0.330824	5.25921359
30		(+) 3	(−) 5	(+) 100	(+) 500	(+) 802.11 b w/out RTS	0.099736	1.04019082
31		(−) 7	(+) 30	(+) 100	(+) 500	(+) 802.11 b w/out RTS	0.16952	3.11871308
32		(+) 3	(+) 30	(+) 100	(+) 500	(+) 802.11 b w/out RTS	0.07568	4.98781013

Table 3: Example Experimental Data

Experiment	x_1	x_2	y
1	(−) 5	(−) 10	0.75
2	(+) 25	(−) 10	0.25
3	(−) 5	(+) 100	0.40
4	(+) 25	(+) 100	0.15

sign points × 5 runs each). Each simulation experiment was executed for 320 seconds. Formally speaking, our approach is a $2^5 5$ factorial design, which implies there are five factors, each at two levels, and the experiment is repeated five times. In addition to the five aforementioned factors that were measured in this study, several other potentially influential factors were held constant. All nodes have a transmission range of 250 meters. The traffic sources were all constant-bit-traffic generators transmitting 512-byte UDP packets at a rate of 2 packets/second. The Location-Aided Routing protocol was used as the routing protocol. The channel propagation model is based on the free-space model with a channel capacity of 2Mbps. The random waypoint mobility model is used to model mobility with a pause-time of 25 seconds.

Step 6: Computing the main and interactive effects. Recall that we are interested in analyzing the main and interactive effects that factors have on specific response metrics. For clarity, we illustrate a simple approach for estimating main and two-factor interaction effeects [9]. Let us consider the 2^2 factorial design shown in Table 3, with factors x_1 and x_2 for which we are interested in quantifying their effect on the response metric y. Notice in experiments 1 and 2 we vary x_1 from its − level to its + level while holding x_2 at its − level. In both cases, we obtain values for the response metric y. Similarly, in experiments 3 and 4 we vary x_1 from its − level to its + level while holding x_2 at its + level. As before, we obtain values for the response metric (y). We can express the functional relationship $y(x_1, x_2)$ using the following effects model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{1}$$

where β_0 is the average response over all simulation runs, β_1 and β_2 represent the main effects of x_1 and x_2, respectively, and β_{12} represents the interactive effect of factors x_1 and x_2, respectively.

Substituting the four response observations y_1, y_2, y_3, and y_4 (one for each design point in a 2^2 design matrix) and the coded values for each factor in Equation 1, we have

$$y_1 = \beta_0 - \beta_A - \beta_B + \beta_{AB} \tag{2}$$

$$y_2 = \beta_0 + \beta_A - \beta_B - \beta_{AB} \tag{3}$$

$$y_3 = \beta_0 - \beta_A + \beta_B - \beta_{AB} \tag{4}$$

$$y_4 = \beta_0 + \beta_A + \beta_B + \beta_{AB} \tag{5}$$

Solving Equations 2, 3, 4, and 5 for β_i's, we have

$$\beta_0 = \frac{1}{2}\left(y_1 + y_2 + y_3 + y_4\right) \tag{6}$$

$$\beta_1 = \frac{1}{2}\left(-y_1 + y_2 - y_3 + y_4\right) \tag{7}$$

$$\beta_2 = \frac{1}{2}\left(-y_1 - y_2 + y_3 + y_4\right) \tag{8}$$

$$\beta_{12} = \frac{1}{2}\left(y_1 - y_2 - y_3 + y_4\right) \tag{9}$$

From these results, we see that the main effect of each factor is actually the difference between two averages:

$$E = \bar{y}_{+} - \bar{y}_{-} \tag{10}$$

where $\bar{y}_{+}$ is the average response when the factor is at its *high* level and $\bar{y}_{-}$ is the average response when the variable is at its *low* level. Furthermore, the interactive effect is the average change in the response metric when the two factors are at the same level (+ or −) and when they are at different levels. It is important to note that all responses for each experimental design point is used to determine all main and joint effects [3].

4. DATA ANALYSIS

In this section, we discuss the results of our statistical DOE, along with an analysis of these results. Specifically, we shall first provide an intuitive and visual illustration regarding the impact of the factors on performance. We the quantify this intuition by way of statistical analysis. For the discussion which follows, the reader may find it helpful to refer to the design matrix shown in Table 2.

4.1 Preliminary Insights

A scatterplot can be used to visualize performance changes as the factor levels are changed. Each value along the x-axis corresponds to a design point (or simulation scenario) as shown in Table 2. The y-axis is the performance metric under consideration, and each point on the graph is the average of $r = 5$ simulations for that particular design point.

Upon inspection of the scatterplots in Figures 1(a) and 1(b), it is important to note that the individual points in each of the scatterplots reflect a change in the *average number of neighbors* factor from its − to its + level (that is, from 7, or strongly-connected, to 3, or weakly-connected). Similarly, point-pairs 1-2/3-4, 5-6/7-8, and so on, reflect a change in the *average node speed* from its − to its + level (i.e., from 5 m/sec to 30 m/sec). This observable pattern can help the researcher determine whether or not particular effects are present between factors and performance metrics.

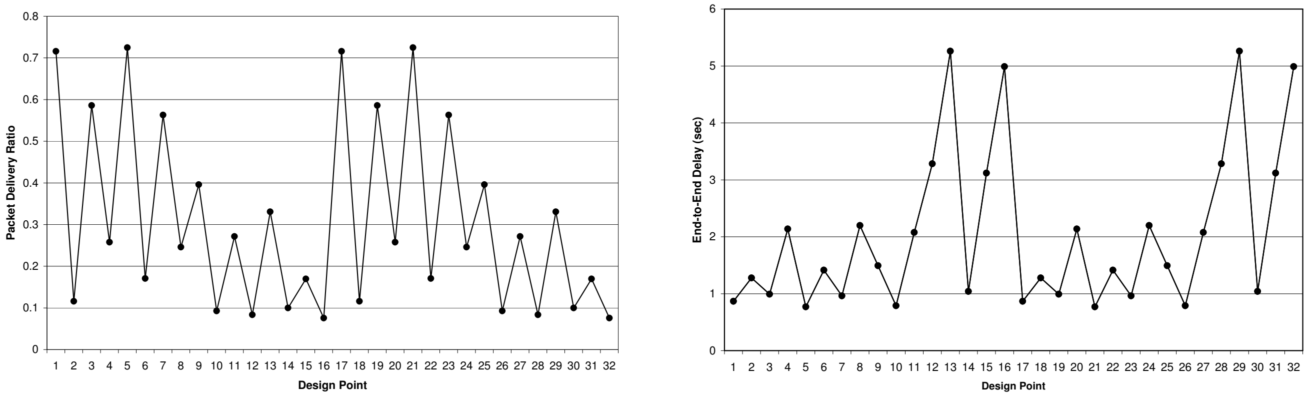
Before we examine the two scatterplots in detail, it is useful to first glean some preliminary insights into what these graphs tell the researcher. The most apparent element when contrasting the two scatterplots is that when end-to-end delay is small, the packet delivery ratio is large (see run numbers 1 through 9 and run numbers 17 through 25 in Figures 1(a) and 1(b)). Conversely, we observe that the packet delivery ratio is small when end-to-end delay is large (see run numbers 10 through 16 and run numbers 26 through 32 in the same two graphs). These observations are reasonable because a smaller end-to-end delay implies that: (1) a greater number of packets are being received by the receiver per unit time when there is very little end-to-end delay; and (2) a smaller number of packets are being received by the receiver per unit time when the end-to-end delay is large.

Packet Delivery Ratio. Figure 1(a) illustrates the average packet delivery ratio for the 32 experimental runs. Observe that the same pattern occurs twice. Specifically, experimental run numbers 17 through 32 exhibit the same general behavior as that of experimental run numbers 1 through 16. The "shift" at run 17 reflects the change in the MAC layer protocol from 802.11b with RTS (− level) to 802.11b without RTS (+ level). By inspection, we may infer that, regarding packet delivery ratio at least, the presence of RTS–or lack thereof–seems to have little or no effect. Next, we observe how the behavior of the packet delivery ratio changes, beginning at the 9th and 25th experimental runs. These are the run numbers at which the number of nodes switches from 100 to 500. Of course, the number of nodes switches from 500 to 100 at run number 17. As can be seen from Figure 1(a), there is a decrease in the variation of average packet delivery ratio as the number of nodes increases. Continuing with our analysis, it appears that varying the traffic load from 10% to 20% has no effect on average packet delivery ratio, relative to the overall number of nodes. A similar observation is made regarding node speed, where there seems to be minimal change in behavior. Finally, when varying the number of neighbors from 7 to 3, the impact on packet delivery ratio is somewhat striking.

End-to-end Delay. Figure 1(b) illustrates the average end-to-end delay for the 32 experimental runs. Here we see a pattern of repetition that resembles that which was discussed for Figure 1(a). As before, experimental run numbers 17 through 32 exhibit the same general behavior as that of experimental run numbers 1 through 16. Again, the "shift" at run 17 reflects the change in the MAC layer protocol from 802.11b with RTS (− level) to 802.11b without RTS (+ level). As with packet delivery ratio, we may infer that, regarding end-to-end delay, the presence of RTS versus its non-presence appears to have little effect. Next, we observe that the behavior of end-to-end delay changes, beginning at the 10th and 26th experimental runs. Given that the number of nodes switches from 100 to 500 at experimental run numbers 9 and 25, there seems to be a slight delay before the effect of this change actually impacts average end-to-end delay. As before, the number of nodes switches from 500 to 100 at run number 17. As can be seen in the figure, there is a substantial "spike" in the variation of average end-to-end delay as the number of nodes increases. Continuing with our analysis, it seems that varying the traffic load from 10% to 20% has minimal impact on average end-to-end delay, relative to the overall number of nodes. The impact on end-to-end delay from varying the node speed between 5 meters/sec and 30 meters/sec appears to be rather substantial, especially as the speed is increased. Finally, as with the packet delivery ratio, the impact on end-to-end delay appears to be very prominent when varying the number of neighbors from 7 to 3.

4.2 Main and Interaction Effects.

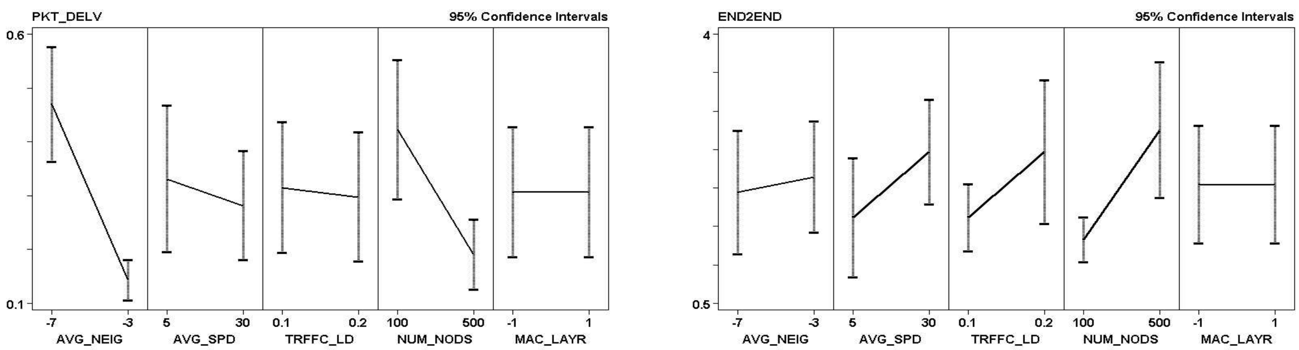
A main effects plot can be used to visualize performance changes as each individual factor level is changed. Each value along the x-axis corresponds to a − or + level for a particular factor as shown in Table 2. The y-axis is the performance metric under consideration, and the line shifts connecting the two points illustrate the average main effect on the performance metric when varying a factor from its − level to its + level. The slope of the line shift for a performance metric by varying a particular factor from its − level to its + level indicates the degree to which the particular factor has a main effect on the performance metric.



(a) Packet Delivery Ratio

(b) End-to-end Delay

Figure 1: Scatterplots can be used to visualize performance changes as the factor levels are varied.



(a) Main effects - Packet delivery ratio

(b) Main effects - End-to-end delay

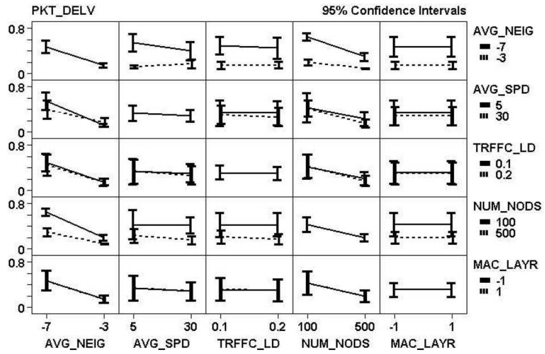
Figure 2: Main Effects

In short, the greater the slope of a line shift, the greater the average main effect upon the performance metric by the particular factor. If a line shift exhibits a small slope (or, for that matter, no slope), then the average main effect upon the performance metric by the particular factor is negligible (or, in the case of no slope, is nonexistent). It is important to keep in mind that the insights gleaned from main effects plots are only for the range of values used for the − and + levels of the factors under consideration.

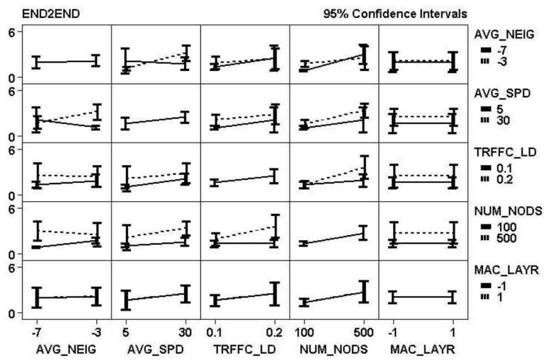
As can be seen in Figure 2(a), the main effect on *packet delivery ratio* by varying *average number of neighbors*, *average speed*, *traffic load*, and *number of nodes* from their − levels to their + levels is apparent. Moreover, it appears from Figure 2(a) that both *average number of neighbors* and *number of nodes* markedly impact the packet delivery ratio, whereas the *MAC layer* has a negligible impact on packet delivery ratio. For example, we see that the packet delivery ratio decreases from roughly 0.4 to roughly 0.2 when the number of nodes is varied from 100 (its − level) to 500 (its + level). In contrast, the packet delivery ratio remains at around 0.3 when varying the MAC layer protocol from 802.11b w/RTS (its − level) to 802.11b w/out RTS (its + level).

Figure 2(b) suggests that the main effects of all factors, except for the *MAC layer*, impact the *end-to-end delay*. For example, we see that the end-to-end delay increases from roughly 1.5 seconds to roughly 2.5 seconds when the average node speed is varied from 5 meters/second (its − level) to 30 meters/second (its + level). In contrast, the end-to-end delay remains at around 2 seconds when varying the MAC layer protocol from 802.11b w/RTS (its − level) to 802.11b w/out RTS (its + level).

Comparing Figures 2(a) and 2(b), we observe that as the average neighbors is varied from “strongly-connected” to “weakly-connected” (that is, when the number of neighbor nodes changes from 7 to 3), the main effect upon packet delivery ratio is such that it is dramatically decreased, with a corresponding slight increase in end-to-end delay. This is likely due to the reduction in the availability of links, since there are fewer neighbor nodes. We observe similar main effects phenomena when varying the average speed, traffic load, and number of node factors from their “−” levels to their respective “+” levels. A particularly significant main effect results from varying the number of nodes from 100 to 500, whereby the packet delivery ratio is drastically reduced



(a) Two-way factor interaction effects - Packet delivery ratio



(b) Two-way factor interaction effects - End-to-end delay

Figure 3: Two-Way Interaction Effects

and end-to-end delay increases substantially. A probable explanation is that the greater number of nodes also leads to increased network traffic, which results in much greater contention of the channel among the nodes in the network. A final point of interest is the fact that the MAC layer protocol has virtually no effect on either performance metric.

Having examined the apparent main effects of each of the factors on the response metrics, we next turn our attention to *interaction effects*, which are those combinational effects that two factors have on the two response metrics. Thus, two-way factor interaction effects plots can be used to visualize the performance changes that result from the combined varying of two factors from their $-$ levels to their $+$ levels. This is particularly important, since such two-way factor interactions are not apparent when using the traditional OFAT approach. Note that parallel lines suggest a lack of factor interaction, whereas non-parallel lines suggest the presence of two-way factor interactions.

Figure 3(a) shows the two-way factor interactions on the average *packet delivery ratio* metric by varying from low to high levels for each factor. From Figure 3(a), we see that the following two-way factor interactions have a notable impact on the *packet delivery ratio*: (1) average number of neighbors and average node speed; (2)average number of neighbors and number of nodes; (3)average node speed and traffic load; and (4) number of nodes and average node speed.

Figure 3(b) shows the two-way factor interactions on the *end-to-end delay* response metric by varying from low to high levels for each of the five factors. The following two-way factor interactions appear to have notable impact on the *end-to-end delay*: (1) average number of neighbors and average node speed; (2) average number of neighbors and traffic load; (3) average number of neighbors and number of nodes; (4) average node speed and traffic load; (5) average node speed and number of nodes; and (6) average traffic load and number of nodes.

These visual observations of two-way factor interactions intuitively correspond with the aforementioned main effects. Similar to what we observed in the main effects graphs, the MAC layer protocol appears to have no apparent two-way factor interaction effects. These observed results are important for researchers when considering new protocol designs, since it is obvious that varying single factors may lead to undesirable performance results. However, an awareness of and knowledge about two-way factor interactions may allow researchers to exploit these interactions in such a way that desirable performance results may be realized.

4.3 Quantifying the Main and Joint Effects

Scatter plots and effects plots offer a graphical and intuitive way of inferring whether main and interactive effects exists. Such "evidence" is not sufficient to draw definitive conclusions regarding factors and their impact on the response metrics. We must go one step further and quantify these effects using statistical analysis. Using a simplified method called the "sign-table" method, which is based on the mathematical properties discussed in Section 3.2, we compute the main and two-way interaction effects for each factor. Performing an analysis of variance (ANOVA) allows us to determine the statistical significance of the main and two-way interaction effects.

Table 4 shows the effect estimate and the allocation of variation for each factor and two-way interaction. The allocation of variation indicates the percentage of response variation contributed to a specific factor or two-way interaction. We see that certain factors account for a large percentage of the performance change. For example, we see in Table 4 that average neighbors and number of nodes together account for almost 85% of the performance change in packet delivery ratio. A similar observation may be made for end-to-end delay, where the average speed and number of nodes factors, as well as the average neighbors and number of nodes two-way interaction, together account for approximately 70% of the performance change.

As shown in Table 4, each factor and two-way interaction has an "estimate" associated with it. This estimate quantifies the change in the performance metric when varying the factor (or two-way interaction) from its "$-$" level to its "$+$" level. For example, we see that the estimate for average neighbors is -0.3269 with respect to packet delivery ratio. Since varying the average neighbors factor from its "$-$" level to its "$+$" level is a *two-unit* change (that is, moving from -1 to $+1$), we take one-half the value of its estimate, -0.3269, which is -0.163452, and say this is the expected change in packet delivery ratio when average neighbors changes by one unit. Table 4 also highlights those factors, as well as the two-way factor interactions, that are statistically significant for the prediction of *end-to-end delay*. Here we see the following individual factors that are statistically significant per their impact on end-to-end delay: *average number of neighbors*, *average node speed*, *traffic load*, and *number of nodes*. The two-way factor interactions that are statistically significant include: *average number of neighbors* and *average node speed*; *average number of neighbors* and *number of nodes*; and *traffic load* and *number of nodes*.

5. EMPIRICAL MODELS

We are now ready to build empirical models that are based on the data we have collected. The statistical DOE approach is what facilitates this. In our preliminary work, we build first-order regression models, which are useful for prediction of response (performance) metrics over the range of scenarios examined. The general *multiple linear regression model* with k regressor variables is of the form

$$y = \Theta_0 + \Theta_1 x_1 + \Theta_2 x_2 + \ldots + \Theta_k x_k + \epsilon \tag{11}$$

The parameter Θ_j represents the expected change in response y per unit change in x_j when all the remaining independent variables x_j $(x \neq j)$ are held constant. (Note that ϵ is an error term.) Estimates of the regressors Θ_j are determined using least-squares. Determining the regression constants in a factorial design is the critical part in developing empirical models. This is straightforward using statistical DOE since $\Theta_j = \frac{1}{2}\beta_j$ (the effects estimates; see Table 4). The reason for this is that the effect estimate is based on a two-unit change (that is, from -1 to $+1$), while the regression estimate is based on a one-unit change. Thus, we can derive a first regression model comprising only the significant factors and two-way interactions as follows:

$$\begin{aligned} y_{pdr} \;=\; & 0.306176 - 0.163452 x_1 - 0.024623 x_2 \qquad (12) \\ & - 0.00865 x_3 - 0.11626 x_4 + 0.047578 x_1 x_2 \\ & + 0.013931 x_1 x_3 + 0.061414 x_1 x_4 \\ & - 0.015257 x_2 x_4 - 0.012326 x_3 x_4 \end{aligned}$$

where x_1 = avg neighbors; x_2 = avg node speed; x_3 = traffic load; x_4 = no. of nodes; and and x_5 = MAC layer.

Equation 12 is interpreted as follows. The mean packet delivery ratio is 0.306176; the effect of average neighbors is -0.163452, which is the expected change in the packet delivery ratio per unit change in average neighbors (that is, when average neighbors is varied from its "$-$" level to its "+" level); the effect of average node speed is -0.024623; the effect of traffic load is -0.00865; the effect of number of nodes is -0.11626; the interaction between average neighbors and average node speed is 0.047578; the interaction between average neighbors and traffic load is 0.013931; the interaction between average neighbors and number of nodes is 0.061414; the interaction between average node speed and number of nodes is -0.015257; and the interaction between traffic load and number of nodes is -0.012326. As already discussed, the MAC layer has no effect on the packet delivery ratio, which is reflected by the absence of the regressor variable x_5 in Equation 12.

Table 5 provides fit statistics for packet delivery ratio. We see that Table 5 contains two columns: "Master Model" and "Predictive Model." The "Master Model" values include *all* the factors and factor interactions, whereas the data in the "Predictive Model" column includes *only* those factors and factor interactions that are statistically significant. Our discussion of the fit statistics will focus on the data in the "Predictive Model" column, since it is precisely a prediction model we seek to develop.

Table 5: Fit statistics for packet delivery ratio

	Master Model	Predictive Model
Mean	0.306176	0.306176
R-square	99.32%	99.14%
Adj. R-square	98.68%	98.78%
RMSE	0.025523	0.024537
CV	8.336023	8.014109

The fit statistics for packet delivery ratio may be interpreted as follows. The mean packet delivery ratio is 0.306176, which we have already seen in the regression model shown in Equation 12; the quantity "R-square" is 99.14%, which is the proportion of total variability explained by the model, where $0 \leq R^2 \leq 1$, with larger values being more desirable; the quantity "Adj. R-square," or *adjusted* R^2, is 98.78%, and is a variation of the R^2 statistic whose value decreases as more factors are included within the model; the RMSE, or *root mean square error*, is 0.024537, which is determined by calculating the deviations of points from their true position, summing up the measurements, and then taking the square root of the sum, smaller values being more desirable; and CV, or *coefficient of variation*, a measure of the precision or relative dispersion, which is 8.014109, and is calculated as the standard deviation divided by the mean.

Table 6 is an ANOVA (analysis of variance) table for packet delivery ratio, which is a useful tool for identifying main and interaction effects of factors that are statistically significant. The sum of squares (SS) is the variation; the degrees of freedom (df) is equal to 1 for each factor and factor interaction; the mean square (MS) is the variance, or SS/df; F is the F-ratio, which is MS/Error; and the P-value, which we have discussed earlier. The P-value is of particular interest to us, since it serves as an indicator of "statistical significance," which indicates the degree to which the value of a factor is "true." The greater the value of a P-value, the greater confidence we can have in its reliability. (It is also worth noting that the P-value is the probability that the computed F-statistic is greater than the F-value, usually found in an F table of most statistics books.)

Factors and factor interactions for which the P-value (see Table 6) is small ($P < 0.05$) are considered significant and should therefore be included in the prediction or regression model. The P-value is an indicator of "statistical significance," which indicates the degree to which the value of a factor is "true." From Table 6 we can see that *average number of neighbors*, *average node speed*, *traffic load*, and *number of nodes* are those factors which are statistically significant in terms of their effect on the *packet delivery ratio*. We see similar statistically-significant effects with the following two-way factor interactions: *average number of neighbors* and *average node speed*; *average number of neighbors* and *traffic load*; *average number of neighbors* and *number of nodes*; *average node speed* and *number of nodes*; and *traffic load* and *number of nodes*.

Based upon our analysis of both the fit statistics and ANOVA table for packet delivery ratio, we conclude that the regression equation for packet delivery ratio is an ac-

Table 4: Effects Table

	Packet Delivery Ratio		E2E Delay	
Effect	Estimate	Allocation of Variation	Estimate	Allocation of Variation
AVG_NEIG	−0.3269	55.727	−0.12728	3.031
AVG_SPD	−0.049245	1.265	−0.34479	22.171
TRFFC_LD	−0.017301	0.156	−0.12024	2.696
NUM_NODS	−0.23252	28.193	-0.31514	18.522
MAC_LAYR	5.20417E−18	0.000	−6.245E−17	0.000
AVG_NEIG × AVG_SPD	0.095157	4.722	−0.22768	9.668
AVG_NEIG × TRFFC_LD	0.027861	0.401	−0.0042055	0.00329
AVG_NEIG × NUM_NODS	0.12283	7.867	0.39482	29.072
AVG_NEIG × MAC_LAYR	6.07153E−18	0.000	0	0.000
AVG_SPD × TRFFC_LD	−0.018785	0.184	0.058292	0.634
AVG_SPD × NUM_NODS	−0.030515	0.486	−0.1002	1.872
AVG_SPD × MAC_LAYR	5.20417E−18	0.000	-1.3878E−17	0.000
TRFFC_LD × NUM_NODS	−0.024651	0.317	−0.14219	3.771
TRFFC_LD × MAC_LAYR	−8.6736E−18	0.000	6.93889E−18	0.000
NUM_NODS × MAC_LAYR	5.20417E−18	0.000	4.85723E−17	0.000

Table 6: ANOVA - Packet Delivery Ratio

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	P Value (Pr > F)
AVG_NEIG	1	0.854925	0.854925	1312.409	0.0001
AVG_SPD	1	0.019401	0.019401	29.78212	0.0001
TRFFC_LD	1	0.002395	0.002395	3.675985	0.073234
NUM_NODS	1	0.432521	0.432521	663.9696	0.0001
MAC_LAYR	1	2.17E−34	2.17E−34	3.33E−31	1
AVG_NEIG ∗ AVG_SPD	1	0.072439	0.072439	111.2021	0.0001
AVG_NEIG ∗ TRFFC_LD	1	0.00621	0.00621	19.532893	0.007062
AVG_NEIG ∗ NUM_NODS	1	0.120692	0.120692	185.2759	0.0001
AVG_NEIG ∗ MAC_LAYR	1	2.95E−34	2.95E−345	4.53E−31	1
AVG_SPD ∗ TRFFC_LD	1	0.002823	0.002823	4.333649	0.053781
AVG_SPD ∗ NUM_NODS	1	0.007449	0.007449	11.43558	0.003805
AVG_SPD ∗ MAC_LAYR	1	2.17E−34	2.17E−34	3.33E−31	1
TRFFC_LD ∗ NUM_NODS	1	0.004861	0.004861	7.462776	0.014776
TRFFC_LD ∗ MAC_LAYR	1	6.02E−34	6.02E−34	9.42E−31	1
NUM_NODS ∗ MAC_LAYR	1	2.17E−34	2.17E−34	3.33E−31	1
Model	15	1.523715	0.101581	155.9386	0.0001
Error	16	0.010423	0.000651		
(Lack of fit)					
(Pure Error)					
Total	31	1.534137			

ceptable predictive empirical model, at least for those values that are within the ranges of the factor levels upon which our experimental design is structured.

Following the same strategy and analysis outlined above for packet delivery ratio, we also derive a predictive model for end-to-end delay. From the ANOVA table we see that the following factors and factor interactions are statistically significant: (1) AVG_NEIG, (2) AVG_SPD, (3) TRFFC_LD, (4) NUM_NODS, (5) AVG_NEIG ∗ AVG_SPD, (6) AVG_NEIG ∗ NUM_NODS, and (6) TRFFC_LD ∗ NUM_NODS.

Hence, we derive the empirical model shown in Equation 13 for end-to-end delay. Based upon our analysis of both the fit statistics and ANOVA table for end-to-end delay, we conclude that the regression equation for end-to-end delay is an acceptable predictive empirical model, at least for those values that are within the ranges of the factor levels upon which our experimental design is structured.

$$\begin{aligned} y_{e2e} \quad = \quad & 0.706945 - 0.06364x_1 - 0.172396x_2 \\ & - 0.060119x_3 - 0.15751x_4 - 0.113838x_1x_2 \\ & + 0.19741x_1x_4 - 0.071095x_3x_4 \end{aligned} \tag{13}$$

where x_1 = avg neighbors; x_2 = avg node speed; x_3 = traffic load; x_4 = no. of nodes; and and x_5 = MAC layer.

Table 8: Fit statistics for end-to-end delay

	Master Model	Predictive Model
Mean	0.706945	0.706945
R-square	91.43%	88.92%
Adj. R-square	83.40%	85.69%
RMSE	0.151566	0.140712
CV	21.43958	19.90431

6. CONCLUSION AND FURTURE WORK

In this paper, our contributions are three-fold. First, we demonstrated how a systematic statistical design of experiments (DOE) approach can be an extremely useful tool for researchers who wish to gain the maximum information with a minimum of experiments. Second, we used a factorial design to quantify the main and two-way interaction effects of five factors (i.e., network density, node mobility, traffic load, network size, and medium access control scheme) on two performance metrics (i.e., packet delivery ratio and end-to-end delay). Third, we show how the main and joint effect estimates can be used developed empirical first-order linear regression models.

The overall goal of our work in this area is to build empirical models that accuarately characterize and predict the performance of mobile ad hoc networking system over a wide range of scenarios. Such models can then be used in the development of an autonomic control system which adjust

Table 7: ANOVA - End-to-end Delay

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	P Value (Pr > F)
AVG_NEIG	1	0.129601	0.129601	5.641656	0.030379
AVG_SPD	1	0.951048	0.951048	41.39988	0.0001
TRFFC_LD	1	0.115658	0.115658	5.034707	0.039346
NUM_NODS	1	0.794518	0.794518	34.58601	0.0001
MAC_LAYR	1	3.12E−32	3.12E−32	1.36E−30	1
AVG_NEIG ∗ AVG_SPD	1	0.414691	0.414691	18.05181	0.000613
AVG_NEIG ∗ TRFFC_LD	1	0.000141	0.000141	0.006159	0.938419
AVG_NEIG ∗ NUM_NODS	1	1.247063	1.247063	54.28567	0.0001
AVG_NEIG ∗ MAC_LAYR	1	0	0	0	1
AVG_SPD ∗ TRFFC_LD	1	0.027184	0.027184	1.183327	0.292795
AVG_SPD ∗ NUM_NODS	1	0.080319	0.080319	3.496352	0.079916
AVG_SPD ∗ MAC_LAYR	1	1.54E−33	1.54E−33	6.71E−32	1
TRFFC_LD ∗ NUM_NODS	1	0.161742	0.161742	7.040762	0.017347
TRFFC_LD ∗ MAC_LAYR	1	3.85E−34	3.85E−34	1.68E−32	1
NUM_NODS ∗ MAC_LAYR	1	1.89E−32	1.89E−32	8.22E−31	1
Model	15	3.921966	0.261464	11.38176	0.0001
Error	16	0.367556	0.022972		
(Lack of fit)					
(Pure Error)					
Total	31	4.289522			

system parameters and protocols selections based on current and predicted performance results. As such, the analysis presented in this paper is only a preliminary step in the process. Future work will include the anslysis of a larger parameter space and the development of second order and non-linear models, which will be based on both regression and response surface modeling techniques as well as neural network models. Model accuracy (i.e., the ability of the models to predict performance) will be verified by comparing the model results with both simulation and testbed results.

7. REFERENCES

[1] C. Barrett, M. Drozda, A. Marathe, and M. V. Marathe. Characterizing the interaction between routing and mac protocols in ad-hoc networks. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking and computing*, pages 92–103, 2002.

[2] C. L. Barrett, M. Drozda, A. Marathe, and M. V. Marathe. Analyzing interaction between network protocols, topology and traffic in wireless radio networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2003.

[3] G. E. P. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. John Wiley and Sons, Inc., New York, NY, 1978.

[4] M. Conti, G. Maselli, G. Turi, and S. Giordana. Cross-layering in mobile ad hoc network design. *IEEE Computer*, 37(2):48–51, February 2004.

[5] S. Corson and J. Macker. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations. *Internet-Draft*, Oct. 1998.

[6] A. Ephremides. Energy concerns in wireless networks. *IEEE Wireless Communications*, August 2002.

[7] Z. J. Haas. Panel report on ad-hoc networks. *Mobile Computing and Communications Review*, 2(1), 1997.

[8] G. Holland and N. H. Vaidya. Analysis of TCP performance over mobile ad hoc networks. In *Proceedings of IEEE/ACM MOBICOM*, pages 219–230, August 1999.

[9] R. Jain. *The Art of Computer System Performance and Analysis*. John Wiley and Sons, New York, 1991.

[10] J. P. C. Kleijnen. A user's guide to the brave new world of designing simulation experiments. White paper, Tilburg University, CentER, March 2004.

[11] S. Kumar, V. S. Raghavan, and J. Deng. Medium access control protocols for ad hoc wireless networks: A survey. *Elsevier Ad-Hoc Networks Journal*, 2004.

[12] Law and Kelton. *Simulation, Modeling and Analysis*. McGraw-Hill Higher Education, New York, 2000.

[13] J. Lee, S. Singh, and Y. Roh. Interlayer interactions and performance in wireless ad hoc network. Internet-draft, IRTF ANS Working Group, September 2003.

[14] D. C. Montgomery. *Design and Analysis of Experiments*. John Wiley and Sons, New York, 2001.

[15] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser. An analysis of the optimum node density for ad hoc mobile networks. In *Proceedings of the IEEE International Communications Conference*, pages 857–861, Helsinki, Finland, June 2001.

[16] M. W. Subbarao. Ad hoc networking critical features and performance metrics. White paper, National Institutes of Standards and Technology, September 1999.

[17] H. Takagi and L. Kleinrock. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Transactions on Communications*, 32(3), Mar. 1984.

[18] K. K. Vadde and V. R. Syrotiuk. Factor interaction on service delivery in mobile ad hoc networks. *IEEE Journal on Selected Areas in Communications*, 22(7):1335–1346, September 2004.