MANET Simulation Studies: The Incredibles *

Stuart Kurkowski *skurkows@mines.edu*
Tracy Camp *tcamp@mines.edu*
Michael Colagrosso *mcolagro@mines.edu*
MCS Department, Colorado School of Mines, Golden, Colorado, USA

Simulation is the research tool of choice for a majority of the mobile ad hoc network (MANET) community. However, while the use of simulation has increased, the credibility of the simulation results has decreased. To determine the state of MANET simulation studies, we surveyed the 2000-2005 proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc). From our survey, we found significant shortfalls. We present the results of our survey in this paper. We then summarize common simulation study pitfalls found in our survey. Finally, we discuss the tools available that aid the development of rigorous simulation studies. We offer these results to the community with the hope of improving the credibility of MANET simulation-based studies.

I. Introduction

Mobile Ad Hoc Networks (MANETs) are wireless mobile nodes that cooperatively form a network without infrastructure. Because there is no coordination or configuration prior to setup of a MANET, there are several challenges. These challenges include routing packets in an environment where the topology is changing frequently, wireless communications issues, and resource issues such as limited power and storage. The leading way to research solutions to these difficult MANET challenges is simulation.

In this paper, we consider the current state of MANET simulation studies published in a premiere conference for the MANET community, i.e., the Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) from 2000-2005 [9]. The results, unfortunately, are discouraging; in general, results published on MANET simulation studies lack believability. There are several factors involved in conducting trustworthy simulation-based research. For our study we focused on the following four areas of credibility in research.

1. Repeatable: A fellow researcher should be able to repeat the results for his/her own satisfaction, future reviews, or further development.
2. Unbiased: The results must not be specific to the scenario used in the experiment.
3. Rigorous: The scenarios and conditions used to test the experiment must truly exercise the aspect of MANETs being studied.
4. Statistically sound: The execution and analysis of the experiment must be based on mathematical principles.

The remainder of the paper will focus on the current state of MANET simulations, our survey results, common pitfalls to avoid, and tools to aid the researcher in conducting simulation studies. The goal of this paper is to raise awareness on the lack of reliability of MANET simulation-based studies. We present our survey results and identify common issues and pitfalls as a starting point for improvement.

I.A. The Current State of MANET Simulation Studies

We conducted a survey of MANET research published in MobiHoc [9]; we only included the full papers in our survey, not the poster papers. Simulation is an often used tool to analyze MANETs; 114 out of the 151 MobiHoc papers published (75.5%) used simulation to test their research.

There are many discrete-event network simulators available for the MANET community [35]. Unfortunately, 34 of the 114 published MobiHoc simulation papers (29.8%) did not identify the simulator used in the research. Figure 1 shows the simulator usage results of the MobiHoc authors that did identify the simulator used. Network Simulator-2 (NS-2) [34] is the most used simulator in MANET research; 35 of the 80 simulation papers that state the simulator used in the simulation study used NS-2 (43.8%).

When the simulator used is not specified within a published paper, the repeatability of the simulation study is directly compromised. The most direct way to make a research project repeatable is to make the

*This work supported in part by NSF Grants ANI-0208352 and ANI-0240558. Research Group's URL is *http://toilers.mines.edu.*

code and configuration files from the simulation study available to the community; unfortunately, in our survey, no paper made a statement about code availability. In addition, the researcher must identify the simulator and version, the operating system, and all variable settings. Repeatability is also based on the scenarios evaluated, the techniques used to avoid initialization bias (influence of empty queues, etc., at the start), and the techniques used to analyze the results. Thus, a published paper must discuss or reference all of these details to meet the repeatability criteria.

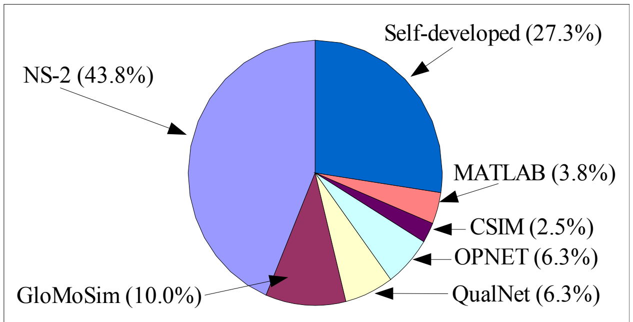

Figure 1: Simulator usage from our MobiHoc survey.

To be an unbiased study, a project must address initialization bias, random number issues, and use a variety of scenarios. The only time to use a single scenario is to prove a limitation or counter a generalization. To be a rigorous study, factors such as node density, node footprint, coverage, speed, and transmission range must be set to exercise the protocol under test. For example, a study that uses scenarios with average hop counts, between source and destination, below two are only testing neighbor communication and not true routing. Finally, to be a statistically sound study, a project must account for initialization bias, execute a number of simulation iterations, provide the confidence levels that exist in the results, and list any statistical assumptions made. In this paper we use the results of our MobiHoc survey to raise awareness of the low percentage of MANET research efforts satisfying these requirements.

I.B. Survey Motivation

The authors of [29] completed a similar evaluation of network simulation studies in 1999. However, because the first MobiHoc conference was in 2000, this previous evaluation of simulation studies was unable to include simulations studies published in the MobiHoc conference. In addition, unlike our paper, the evaluation of simulation studies from 1999 was on network simulations in general, not on MANETs in specific. Because our research is focused on the specific niche of network simulations with mobility, we completed a survey on the state of MANET simulations published in all of the previous MobiHoc proceedings. We found that, although it has been six years since the previous survey study, network simulation studies (at least in the MANET community) have not improved and, in some cases, have deteriorated even further.

As an example where the reliability of simulation studies has not improved, consider the simulation type (i.e., terminating or steady-state) used in a simulation study[1]. In [28], 1690 of 2200 simulation papers (approx. 77%) did not state the type of simulation. In our MobiHoc survey, 66 of the 114 simulation papers (57.9%) did not mention the type of simulation used in the study. As an example where the credibility of simulation studies has deteriorated, consider the pseudo random number generator (PRNG) used in a simulation study. In [28], approximately 650 of the 2200 (approx. 30%) papers stated which PRNG was used in the research. In our MobiHoc survey, not a single paper mentions the PRNG used.

As the MANET community moves forward toward implementation, it is imperative to have reliable simulation research and researchers addressing the design of experiments (DOE) used in their studies [4, 24]. While DOE should be used to conduct the overall study, we leave the DOE details to the DOE community and focus on issues specific to MANET research in this paper.

The remainder of this paper is organized as follows. In Section II, we provide detailed descriptions and results from our survey of the published papers in the 2000-2005 proceedings of the MobiHoc conference. We then document a list of pitfalls that exist in simulation-based MANET studies in Section III. The list was developed from our survey of MobiHoc papers and our own experiences in MANET simulations. Section IV introduces tools researchers can use to conduct credible simulation based studies. Our goal is to raise awareness of the issues and to introduce tools to aid MANET researchers in conducting and reporting credible simulation results.

II. Survey Results

To evaluate the current state of reliability in MANET research we surveyed the published papers of Mobi-

[1] Terminating simulations have a finite end time; steady-state simulations are not time specific, and answer the question of long term performance [18].

http://folk.uio.no/paalee/research.htm

Table 1: Survey results for 151 published papers in ACM's MobiHoc conference, 2000-2005.

Simulator and Environment		
Totals	**Percentage**	**Description**
114 of 151	75.5%	Used simulation in the research.
0 of 114	0.0%	Stated the code was available to others.
80 of 114	70.2%	Stated which simulator was used.
35 of 80	43.8%	Used the NS-2 simulator.
8 of 80	10.0%	Used the GloMoSim simulator.
5 of 80	6.3%	Used the QualNet simulator.
5 of 80	6.3%	Used the OPNET simulator.
3 of 80	3.8%	Used MATLAB/Mathematica.
2 of 80	2.5%	Used the CSIM simulator.
22 of 80	27.3%	Used self-developed or custom simulators.
7 of 58	12.1%	Stated which version of the public simulator was used.
3 of 114	2.6%	Stated which operating system was used.
8 of 114	7.0%	Addressed initialization bias.
48 of 114	42.1%	Addressed the type of simulation.
0 of 114	0%	Addressed the PRNG used.
Simulation Input Parameters		
Totals	**Percentage**	**Description**
109 of 114	95.6%	Conducted MANET protocol simulation studies.
62 of 109	56.9%	Stated the number of nodes used in the study.
58 of 109	53.2%	Stated the size of the simulation area.
62 of 109	56.9%	Stated the transmission range.
49 of 109	45.0%	Stated the simulation duration.
41 of 109	37.5%	Stated the traffic send rate.
31 of 109	28.4%	Stated the traffic type (e.g., CBR, etc.)
39 of 109	35.8%	Stated the number of simulation runs (iterations).
42 of 109	38.5%	Used mobility in the study.
34 of 42	81.0%	Stated the mean speed of the nodes.
26 of 42	61.9%	Stated the speed variance about the mean.
21 of 42	50.0%	Stated the mean pause time of the nodes.
16 of 42	38.1%	Stated the pause time variance about the mean.
38 of 42	90.5%	Stated which mobility model was used.
25 of 38	65.8%	Used the random waypoint mobility model [16].
2 of 25	8.0%	Used the steady-state version of the random waypoint mobility model [26].
2 of 38	5.3%	Used a group mobility model [14, 32].
4 of 38	10.5%	Used a grid/road mobility model (e.g., [7]).
5 of 38	13.2%	Used the random direction mobility model (e.g., [36]).
Plots/Charts/Graphs		
Totals	**Percentage**	**Description**
112 of 114	98.2%	Used plots to illustrate the simulation results.
14 of 112	12.5%	Used confidence intervals on the plots.
100 of 112	89.3%	Had legends on the plots.
84 of 112	75.0%	Had units on the data or labels.

Hoc, a premiere MANET conference. For each paper in the proceedings, we distilled the answers to several simulation study questions. Only the appropriate questions were asked of each paper, e.g., if a paper did not use plots, the detailed plot questions were not surveyed for that paper. Additionally, we reviewed each paper individually avoiding word searches or other means of automatically gathering results; in other words, papers that described the study without using explicit descriptors were counted. For consistency, the same person reviewed all of the papers; to validate the results, we had a second person review all of the papers with a subset of the questions and a third person to correct the few inconsistencies found.

We used the database of survey data to compile the results shown in Table 1, and we discuss some of these results in Section III. Overall, the results in Table 1 indicate trends in the lack of believability in MANET simulation research, even though using MANET simulation research to test performance is prominent; that is, 114 out of the 151 (75.5%) published MobiHoc papers used simulation as the basis for the study. Simulation is a large portion of the research in the MANET community making its lack of believability a concern.

III. Common Simulation Pitfalls

We have developed a list of simulation pitfalls that impact the reliability of a simulation-based study. We have accumulated the list from our own experiences with simulations as well as the experience of others in the field. Pitfalls identified from our survey of MobiHoc papers are also included in the list. Because the pitfalls impact different phases of a simulation-based study, we have grouped the pitfalls into the following categories: simulation setup, simulation execution, output analysis, and publishing.

III.A. Simulation Setup

Simulation setup is the phase of a MANET research effort that is most often skipped or overlooked; and if the setup phase is done improperly, the credibility of the simulation study is flawed from the start. Setup begins with determining the simulation type, validating the model, validating the PRNG, defining variables, and developing scenarios.

III.A.1. Simulation Type

Although selecting the type of simulation appears to be a trivial step, not identifying the type of simulation (terminating vs. steady-state) is a commonly overlooked step for researchers. As mentioned, 66 out of the 114 simulation papers (57.9%) in our MobiHoc survey did not state whether the simulation was terminating or steady-state. We suspect most simulations are steady-state because MANET researchers are typically interested in the long term average behavior of an ad hoc network.

Not determining the simulation type can lead to poorly designed simulations with statistically unsound results. The most common error made by researchers is to execute one type of simulation and report results on the other type of simulation. For example, executing a terminating simulation for a set number of seconds and claiming the results represent the steady-state behavior [28]. This can produce results much different from the steady-state if the simulation terminated well before the statistics converged. The researcher should always determine the type of simulation and measure convergence if it is a steady-state simulation (see Section III.B.2 for more detail). See [22] for an example of a MobiHoc paper identifying the simulation type used in the study.

III.A.2. Model Validation & Verification

After the type of simulation is determined, the simulation model itself must be prepared. As stated in [23] the model must be validated as a baseline to start any experimentation. Many researchers download the NS-2 simulator, compile it, and begin to execute simulations with a model that has not been validated in his or her environment. Additionally, many researchers make changes to NS-2 during the study and these modifications or enhancements need to be validated. Likewise, the protocol that is being evaluated must be verified to ensure it has been coded correctly and operates in accordance with the protocol specifications [2]. Not validating the model or verifying code is a common pitfall [1]. For example, when we upgraded to a new compiler we found that it implemented a broadcast function in one of our protocols differently than before. This difference had an impact on protocol performance. See [41] as an example of MobiHoc authors discussing validation prior to evaluation.

III.A.3. PRNG Validation & Verification

With the computing power available to researchers today and the complexity of the NS-2 model, MANET researchers need to ensure the PRNG is sufficient for his or her study. For example, the NS-2 PRNG does not allow a separate request stream for each dimension (i.e., a unique request stream) that exists in a simulation study. A three dimension example is when a simulation has three different random pieces, such as jitter, noise, and delay. A researcher wants all three of these series (dimensions) to be uniformly distributed with each other and within each stream (e.g., the jitter stream needs to be uniformly distributed). The authors of [19, 28, 29, 30] show that a 2-dimensional request on a PRNG is valid for approximately $8\sqrt[3]{L}$, where L is the cycle length. In NS-2, the cycle length is $2^{31}-1$, which means that only (approximately) 10,000 numbers are available in a 2-dimensional simulation study. Thus, [30] estimates that the NS-2 PRNG is only valid for several thousand numbers before the potential non-uniformity of numbers or the cycling of numbers. This cycling time occurrence is obviously dependent on the number of PRNG calls made during a simulation, but the study in [30] found most network simulations spent as much as 50% of the CPU cycles generating random numbers. Our testing of PRNG cycling shows cycling impact is minimal because the repeat of numbers does not occur with the simulator in the exact same state as the previous time. However, according to [30], the dimensionality of the numbers is likely to cause a problem in correlation. Thus, before publishing results, a researcher should validate the PRNG to ensure the PRNG did not cause correlation in the results. If the cycle length is an issue with NS-2, Akaroa-2 [11] offers an NS-2 compatible PRNG with a cycle of $2^{191}-1$. The Akaroa-2 [11, 31] PRNG provides several orders of magnitude more numbers and is valid to 82 dimensions.

III.A.4. Variable Definition

NS-2 uses hundreds of configurable variables during an execution in order to meet its general wired and wireless network simulator requirements. For example, there are 538 variables defined in the `ns-default.tcl` file of NS-2.1b7a and there are 674 variables defined in the `ns-default.tcl` file of NS-2.27. The large number of variables makes it difficult to track each variable's default setting. Additionally, an increase in the number of variables between the different NS-2 versions indicates there is a rising number of variables with each new version of NS-2. Our review of the Tcl driver files from our protocols, as well as the examples provided by NS-2, show that many simulation driver files leave key parameters undefined. For example, three out of 12 (25%) of the wireless examples in NS-2 do not define the transmission range of a node. The transmission range is a key variable in MANET performance. If the transmission range default is changed from one NS-2 version to the next, the results of a simulation would be significantly different. The researcher should define all of the variables by using his or her own configuration file or Tcl driver file [4]. See [33] as an example of how to define variables and reference them on a website, providing more detail than can be written in a published paper.

III.A.5. Scenario Development

Table 2 lists the parameters used by the authors who provided the number of nodes, the size of the simulation area, *and* the transmission range of nodes used in the simulations. Only 48 of the 109 MANET protocol simulation papers in our survey of published MobiHoc papers provided all three of these input parameters, detailing 61 simulation scenarios. Table 2 shows the wide range of values in these 61 scenarios. We note that scenario #36 and scenario #37 are the only two scenarios that match; the other scenarios are all unique. The number of nodes in a scenario ranged from 10 nodes to 30,000 nodes. The simulation area ranged from 25 m x 25 m to 5000 m x 5000 m. The transmission ranges varied from 3 m to 1061 m. Table 2 also shows the variety of width and height values, illustrating the different shapes used in MANET simulation scenarios. Additionally, Table 2 reflects that the parameter values are often very specific, e.g., a 1981.7 m squared simulation area. The survey results highlight the wide range of simulation scenarios used to conduct MANET research and the lack of uniform rigorous testing of MANET protocols.

We validated the wide range of input parameters by comparing the derived parameters of each scenario. Table 3 shows a list of the derived parameter definitions and formulas. The derived parameters aggregate multiple input parameters to further characterize a scenario. The derived parameters also provide a common basis for comparison across scenarios. Figure 2 is a scatter plot of all the derived parameters for the 61 sets of input parameters. The plot shows every variable plotted against all the others. For example, the upper right plot is simulation area versus neighbor count with edge effect. The scatter plot reflects the wide range of scenarios and the lack of correlation between parameters.

Figure 2 also shows the lack of independence between parameters, such as node density and node coverage. In addition, the lack of multiple groupings in each plot illustrates that the community is not covering the range of values in a consistent organized manner. For example, if there were benchmark scenarios for small, medium, and large sized simulations, then there would be three groupings of values in each of

Table 2: Input parameters from 61 published scenarios in the proceedings of the MobiHoc conference, 2000-2005, sorted by number of nodes.

No.	# Nodes	Area (m x m)	Range (m)
1	10	1000 x 1000	100
2	20	350 x 350	100
3	20	1000 x 750	250
4	24	800 x 1200	250
5	25	200 x 200	100
6	25	900 x 900	250
7	30	350 x 350	100
8	36	3000 x 3000	1061
9	40	350 x 350	100
10	40	900 x 900	250
11	40	5000 x 5000	250
12	50	40 x 40	10
13	50	350 x 350	100
14	50	500 x 500	100
15	50	1500 x 300	250
16	50	1500 x 300	275
17	50	1000 x 1000	250
18	50	1000 x 1000	100
19	60	350 x 350	100
20	70	25 x 25	10
21	70	350 x 350	100
22	80	350 x 350	100
23	90	350 x 350	100
24	100	100 x 100	20
25	100	350 x 350	100
26	100	300 x 1500	250
27	100	400 x 400	100
28	100	1200 x 1200	250
29	100	500 x 500	100
30	100	575 x 575	250
31	100	575 x 575	125
32	100	650 x 650	67
33	100	1000 x 1000	250
34	100	1000 x 1000	150
35	100	1000 x 1000	50
36	100	1000 x 1000	100
37	100	1000 x 1000	100
38	100	2200 x 600	275
39	100	2000 x 600	250
40	100	150 x 1500	250
41	100	3000 x 900	250
42	100	1000 x 1000	100
43	110	350 x 350	100
44	120	2500 x 1000	250
45	200	100 x 100	40
46	200	500 x 500	70
47	200	1700 x 1700	250
48	200	1981.7 x 1981.7	250
49	225	100 x 100	20
50	225	600 x 600	100
51	400	100 x 100	20
52	400	800 x 800	100
53	500	3000 x 3000	67
54	600	3000 x 3000	250
55	625	1000 x 1000	100
56	1000	40 x 40	3
57	1000	81.6 x 81.6	300
58	1000	100 x 100	10
59	1000	500 x 500	20
60	10000	600 x 600	35
61	30000	5000 x 5000	100

the simulation area plots. Finally, the extreme values in the derived parameters do not correlate with the extreme input parameters. For example, the highest number of nodes (30,000) is the 6th lowest value for the neighbor count derived parameter. The MANET community lacks consistent rigorous scenarios to validate and test solutions to MANET issues.

As a result of this lack of rigorous scenarios, researchers need to analyze the topologies generated by the mobility model generators and evaluate the impact of the various scenario parameters. There have been several emails on the NS-2 mailing list [12] asking what a valid scenario is for MANET research, but currently there is no single benchmark of MANET scenarios to test a protocol. The MANET community needs a way to characterize simulation scenarios in order to evaluate and compare protocols and performance and ensure protocols are rigorously tested. For example, from Table 2, scenario #8, the simulation area is 3000 m x 3000 m, but the transmission range of 1061 m lowers the average hop count to only 1.67 hops. This hop count means most source and destination pairs are direct neighbors and the rest have only one intermediate node. See Section IV for tools that aid in scenario evaluation and characterization.

III.B. Simulation Execution

Executing the simulation is where a lot of time is spent. Therefore, it is important to conduct the execution portion correctly. We highlight several execution pitfalls we have discovered; these pitfalls impact data output, analysis, and ultimately results.

Table 3: Derived scenario parameter definitions and formulas.

Parameter	Description	Formula
Simulation Area	Square meter area of the topology.	$w \times h$
Node Density	Density of nodes in the simulation area.	$\frac{n}{(w \times h)}$
Node Coverage	Area covered by a node transmission.	$\pi \times r^2$
Footprint	Percentage of the simulation area covered by a node's transmission range	$\frac{(\pi \times r^2)}{(w \times h)} \times 100$
Maximum Path	The maximum linear distance a packet can travel from source to destination.	$\sqrt{(w^2 + h^2)}$
Network Diameter	The minimum number of hops a packet can take along the maximum path from source to destination.	$\frac{\sqrt{(w^2 + h^2)}}{r}$
Neighbor Count	The number of neighbor nodes based on transmission and simulation area. It does not account for the edge of the simulation area.	$\frac{(\pi \times r^2)}{\left(\frac{w \times h}{n}\right)}$
Neighbor Count Edge Effect	The average number of neighbor nodes accounting for the edge of the simulation area reducing the node's coverage. For example, a node in the corner of the simulation area only has neighbors in 25% of its coverage area.	Simulation with n, r, and $(w \times h)$
w = width, h = height r = transmission range, n = # of nodes		

III.B.1. Setting the PRNG Seed

One mistake we have seen in NS-2 based simulation studies concerns not setting the seed of the pseudo random number generator (PRNG) properly. NS-2 uses a default seed of 12345 for each simulation run [27]. Thus, if an NS-2 user does not set the seed, each simulation will produce identical results. Additionally, if the seed is not set or is set poorly, it can negate the independent replication method which is typically used in analysis. Introducing correlation in the replications negates the common statistical analysis techniques and the results. In our MobiHoc survey, none of the 84 simulation papers addressed PRNG issues. The researcher should ensure the seed is set correctly in his or her Tcl driver file and that the NS-2 `Random` class is used for all random variables.

III.B.2. Scenario Initialization

Another pitfall is not initializing the scenario correctly. This pitfall usually occurs from a lack of understanding of the two types of simulation. In terminating simulations, the network is usually started in a certain configuration that represents the start of the simulation window. For example, if the researcher is trying to simulate a protocol's response to a failure event, he or she needs to have the failure as the initialization of his or her analysis. Likewise, most simulations start with empty caches, queues, and tables. The simulation fills the caches, queues, and tables until a steady-state of activity is reached. Determining and reaching the steady-state level of activity is part of the initialization. Data generated prior to reaching steady-state is biased by the initial conditions of the simulation and cannot be used in the analysis. Steady-state simulations require that the researcher address initialization bias [25, 40]. For example, in protocols that maintain neighbor information, the size of the neighbor table should be monitored to determine when the table entries stabilize, because the protocol will perform differently with empty routing tables. Akaroa-2 [11] is a tool that monitors variables during execution to determine steady-state (See Section IV).

Unfortunately, only eight of the 114 simulation papers in our MobiHoc survey (7.0%) addressed initialization bias, and all eight use the unreliable method of arbitrarily deleting data. The arbitrary discard periods ranged from 50 seconds to 1000 seconds. Deleting the first portion of the data collected is not a plausible solution. There needs to be statistical rigor in determining a simulation has truly reached steady-state. The researcher should monitor convergence for the steady-

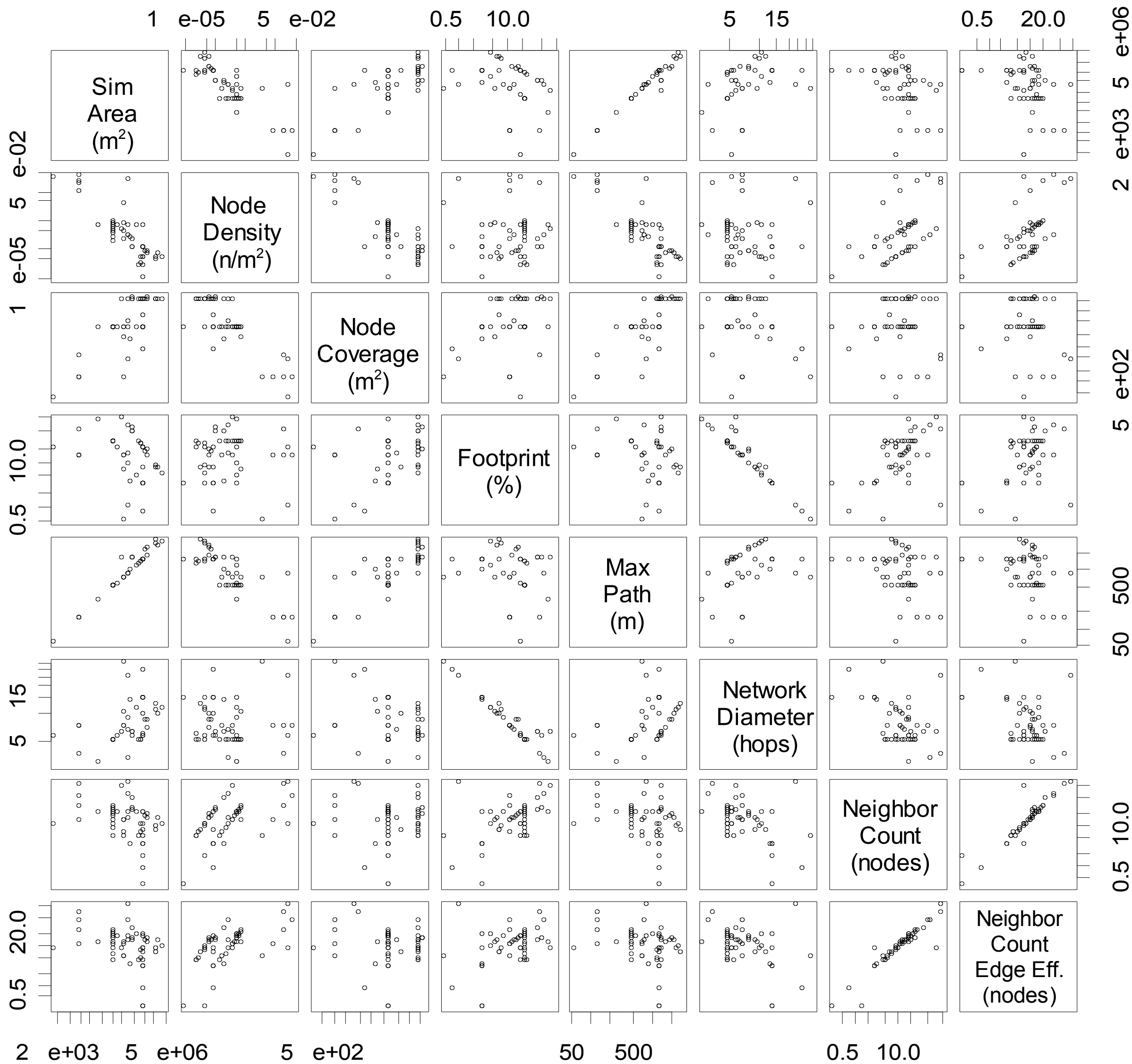

Figure 2: A scatter plot with each of the eight derived scenario parameters plotted against the other derived scenario parameters.

state portions of his or her protocol. For more information on statistically sound methods of addressing initialization bias see [5, 18, 38, 40]. See [8] for an example of a MobiHoc paper that addressed scenario initialization.

III.B.3. Metric Collection

Another area of concern is the metric measurements collected during execution. If the simulation executes properly, but the researcher does not obtain the data he or she needs from the simulation, the simulation is worthless [29]. Appropriate output is especially critical if output has to be correlated. For example, if the researcher is trying to track delivery ratio for data packets and control packets, each type of packet must be identified along with the source and destination to determine the number of each type of packet sent and successfully received. Outputting only the number of packets sent and the number of packets received will not provide the granularity required in the measures. The researcher needs to include output analysis in his or her practice runs of the simulation to ensure the correct metric is being collected. See [20] for an example of a MobiHoc paper describing and defining the statistics used in calculating results.

III.C. Output Analysis

Output analysis is the downfall of many simulation studies. Typically, the preceding steps take longer than planned, which means sufficient time is not provided for output analysis at the end of the schedule. Whether it is the publication deadline, or a thesis defense date, proper analysis is often compromised in the following ways.

III.C.1. Single Set of Data

This pitfall is taking the first set of results from a simulation and accepting the results as "truth". The decision to take the first set is not a plausible way to conduct research. With a single result the probability is high that the single point estimate is not representative of the population statistics. A single execution of a discrete-event simulation is not accounting for the model's innate randomness in the experiment. Executing the simulation once will produce results, maybe even good results [18], however, the single point estimate produced will not give the researcher sufficient confidence in the unknown population mean. The researcher needs to determine the number of runs necessary to produce the confidence levels required for his or her study. In our MobiHoc survey, only 39 of the 109 MANET protocol simulation papers (35.8%) stated the number of simulation runs executed. See [15] for an example of a MobiHoc paper using multiple replications to achieve high confidence and [8] for an example of a MobiHoc paper documenting the number of replications used and how the quantity was chosen.

III.C.2. Statistical Analysis

This pitfall concerns not using the correct statistical formulas with the different forms of output. For example, using the standard formulas for mean and variance without ensuring the data is independent and identically distributed (*iid*). Use of *iid* based formulas on correlated data can reduce reliability by producing biased results. The researcher needs to use batch means or independent replications of the data to ensure *iid* and prevent correlated results [10]. From the survey in [29] 76.5% of the papers did not discuss the statistical methods used in analysis. See [37] for an example of a MobiHoc author that described the analysis and data used to calculate the results.

III.C.3. Confidence Intervals

This pitfall is a culmination of several of the previous analysis issues. Confidence intervals are a tool to provide a range where we think the population mean is located relative to the point estimate [6, 38]. Confidence intervals account for the randomness and varied output from a stochastic simulation. However, in our survey, 98 of the 112 simulation papers using plots (87.5%) did not show confidence intervals on the plots. See [41] for an example of a MobiHoc paper that used confidence intervals.

III.D. Publishing

Table 1 lists all the data from our MobiHoc paper survey. The lack of consistency in publishing simulation-based study results directly impacts the trustworthiness of these studies. In addition, the inconsistency prevents the direct comparison of results, limiting research advancements. The publishing pitfalls prevent the MANET community from taking advantage of new researchers interested in these studies. A new researcher cannot repeat the studies to start his or her own follow-on research.

Publishing is a big part of breaking the "repeatable" criteria for credible research, because much of the simulation study is unknown to the paper reader. As stated earlier, there are 674 variables defined in the `ns-default.tcl` file of NS-2.27. To ensure repeatability the researcher must document the `ns-default.tcl` file used and any changes made to the settings of the variables in the file. When publishing, the authors need to state if the code is available and how to obtain the code. There should be a code statement even if the code's release is restricted by copyright or third party ownership. See [33] as an example of how to properly define variables without using a large portion of the published paper.

At the bottom of Table 1 are publishing specific statistics. Plots of simulation results are common, i.e., 112 of the 114 simulation papers (98.2%) used plots to describe results. However, 12 of the 112 simulation papers with plots (10.7%) did not provide legends or labels on his or her charts. Additionally, 28 of the 112 simulation papers with plots (25.0%) did not provide units for the data being shown. The lack of labels and units can cause readers of these papers to misinterpret or misunderstand the results.

Several of the results in Table 1 are significant inefficiencies in publishing simulation based results. For example, 47 of the 109 MANET protocol simulation papers (43.1%) did not state the transmission range of the nodes. Also, 78 of the 109 MANET protocol simulation papers (71.6%) did not mention the packet traffic type used in the simulation. Although both of these parameters were set to execute the simulation, neither were documented nor referenced in these papers.

A final area of concern in publishing results, one that was not quantified in our survey, is supporting the text with charts and graphs and vice versa. Many papers had charts that were not discussed in the text or the text referenced a chart as supportive, but it was not clear in the chart how it supported the work.

These publishing pitfalls directly impact the credibility of the research conducted in the MANET community. The best simulation based studies can be lost behind a biased, unrepeatable, and unsound document describing the work.

IV. Community Resources

There is some research in developing techniques and processes to aid credible simulation studies. This research is often found in the general simulation community, not the MANET community specifically; however, many groups and authors, such as [1, 3, 13, 39], have outlined steps applicable to MANET research. These methods aid in validation, verification, output analysis, etc. for a simulation based study, and give the overall study more credibility.

Although there has been work on techniques and processes, we have found very few tools that aid researchers in conducting credible simulation studies. Simulation tools are needed to understand the large amount of data produced during network simulations. Tools can analyze the input data as well as aid in validation, verification, initialization, and output analysis. The few tools available today that we are aware include:

- The Akaroa-2 [11] suite, which help a researcher monitor simulation execution to determine steady-state and prevent correlation among multiple replications of a simulation.

- The interactive NS-2 protocol and environment confirmation tool (iNSpect) [17], which visualizes the trace file of an NS-2 simulation. The visualizations can be used for scenario development, model validation, protocol verification, and results analysis.

- The Simulator for Wireless Ad Hoc Networks (SWAN) [21], which enables a researcher to create a virtual environment for conducting experiments with MANETs.

We also note that we are developing a SCenariO characteRizEr for Simulation (SCORES) tool. SCORES will evaluate the rigor with which a scenario tests a MANET protocol by characterizing the scenario.

To aid the community in learning about current and future tools available for use with MANET simulation studies, we have created an on-line list. If you know of a simulation tool that can be used to aid the development of credible simulation studies, please let us know. The current list of tools can be found on our research website at *http://toilers.mines.edu.*.

V. Conclusions

Summarizing the four areas of credibility, we found less than 15% of the published MobiHoc papers are repeatable. It is difficult, if not impossible, to repeat a simulation study when the version of a publicly available simulator is unknown, and only seven of the 58 MobiHoc simulation papers that use a public simulator (12.1%) mention the simulator version used. It is also difficult, if not impossible, to repeat a simulation study when the simulator is self-developed and the code is unavailable. In addition, only eight of the 114 simulation papers (7.0%) addressed initialization bias and none of the 84 simulation papers addressed random number generator issues. Thus, we are concerned that over 90% of the MobiHoc published simulation results may include bias. With regard to compromising statistical soundness, 70 of the 109 MANET protocol simulations papers (64.2%) did not identify the number of simulation iterations used, and 98 of the 112 papers that used plots to present simulation results (87.5%) did not include confidence intervals. Hence, only approximately 12% of the MobiHoc simulation results appear to be based on sound statistical techniques.

MANET simulation-based research is an involved process with plenty of opportunities to compromise the credibility of the study. In this paper, we have identified several pitfalls throughout the simulation lifecycle. Each of the pitfalls discussed in Section III takes away from the goals of making the research repeatable, unbiased, rigorous, and statistically sound. Documenting these pitfalls and sharing knowledge about how to address these common issues will increase the reliability of studies in the MANET community. Our survey of MobiHoc papers showed the current state of MANET research and the lack of consistency, re-enforcing the need for simulation study guidance.

VI. Acknowledgments

We thank Dr. Krzystof Pawlikowski for alerting us to the poor state of simulation based studies in the networking community. Thank you also to Feng Sun, Patrick Milvich, and Tuli Mushell who reviewed the

MobiHoc papers. Finally, we thank the anonymous reviewers for helping us to improve this paper.

References

[1] R. Bagrodia and M. Takai. Position paper on validation of network simulation models. Computer Science Department of University of California, Los Angeles, 1999.

[2] O. Balci. Validation, verification, and testing techniques throughout the lifecycle of a simulation study. In *Proceedings of the 26th Conference on Winter Simulation*, pages 215–220, 1994.

[3] O. Balci. Verification, validation and accreditation of simulation models. In *Proceedings of the 29th Conference on Winter Simulation*, pages 135–141, 1997.

[4] R. Barton. Design of experiments: designing simulation experiments. In *Proceedings of the 33rd Conference on Winter Simulation*, pages 47–52, 2001.

[5] J.-Y. Le Boudec and M. Vojnovic. Perfect simulation and stationarity of a class of mobility models. In *Proceedings of the 38th IEEE Symposium on Simulation*, pages 72–79, 2005.

[6] L. Brakmo and L. Peterson. Experiences with network simulation. In *Proceedings of the Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pages 80–90, 1996.

[7] L. Briesemeister and G. Hommel. Role-based multicast in highly mobile but sparsely connected ad hoc networks. In *Proceedings of the 1st ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 45–50, 2000.

[8] T. Dyer and R. Boppana. A comparison of TCP performance over three routing protocols for mobile ad hoc networks. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 56–66, 2001.

[9] Association for Computing Machinery. The ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc). URL: http://www.sigmobile.org/mobihoc. Page accessed on April 23, 2005.

[10] D. Goldsman and G. Tokol. Output analysis: output analysis procedures for computer simulations. In *Proceedings of the 32nd Conference on Winter Simulation*, pages 39–45, 2000.

[11] Akaroa Research Group. Akaroa2. URL: http://www.tkn.tu-berlin.de/research/ns-2_akaroa-2/ns.html. Page accessed on April 21, 2005.

[12] The VINT Group. Ns-users mailing list. http://mailman.isi.edu/mailman/listinfo/ns-users. Page accessed on April 23, 2005.

[13] J. Heidemann, S. Kumar, and K. Mills. Expanding confidence in network simulations. *IEEE Network*, September/October:58–63, 2001.

[14] X. Hong, M. Gerla, G. Pei, and C. Chiang. A group mobility model for ad hoc wireless networks. In *Proceedings of the ACM International Workshop on Modeling and Simulation of Wireless and Mobile Systems (MSWiM)*, pages 53–60, August 1999.

[15] Y. Hu and D. Johnson. Implicit source routes for on-demand ad hoc network routing. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 1–10, 2001.

[16] D. Johnson and D. Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, pages 153–181. 1996.

[17] S. Kurkowski, T. Camp, and M. Colagrosso. A visualization and animation tool for NS-2 wireless simulations: iNSpect. In *Proceedings of the 13th Annual Meeting of the IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2005.

[18] A. Law and W. Kelton. *Simulation modeling and analysis*. McGraw-Hill, 2000.

[19] P. L'Ecuyer and R. Simard. On the performance of birthday spacing tests for certain families of random number generators. *Mathematics and Computers in Simulation*, pages 131–137, 2001.

[20] S. Lee and C. Kim. Neighbor supporting ad hoc multicast routing protocol. In *Proceedings of the 1st ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 37–44, 2000.

[21] J. Liu, L. F. Perrone, Y. Yuan, and D. Nicol. The simulator for wireless ad hoc networks (SWAN). URL: http://www.eg.bucknell.edu/swan/. Page accessed on August 19, 2005.

[22] Tommaso Melodia, Dario Pompili, Vehbi C. Gungor, and Ian F. Akyildiz. A distributed coordination framework for wireless sensor and actor networks. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, pages 99–110, 2005.

[23] S. Midkiff. Simulation: Methodology. Presentation to ECE590 class. URL: http://fiddle.visc.vt.edu/courses/ece5984-net2/lectures/09_sim.pdf. Page accessed on April 15, 2005.

[24] D. Montgomery. *Design and Analysis of Experiments*. Wiley, 2005.

[25] E. Mota, A. Wolisz, and K. Pawlikowski. Experience with Akaroa: speeding up stochastic simulation using parallel replications. In *Proceedings of the High Performance Computing and Advanced Simulation Technologies Conference*, pages 296–301, 1998.

[26] W. Navidi and T. Camp. Stationary distributions for the random waypoint mobility model. *IEEE Transactions on Mobile Computing*, pages 99–108, 2004.

[27] *The NS manual: formerly known as notes and documentation*, 2003. URL: http://www.isi.edu/nsnam/ns/. Page accessed on April 21, 2005.

[28] K. Pawlikowski. Do not trust all simulation studies of telecommunications networks. *International Conference on Information Networking (ICOIN)*, pages 899–908, 2003.

[29] K. Pawlikowski, J. Jeong, and R. Lee. On credibility of simulation studies of telecommunication networks. *IEEE Communications Magazine*, pages 132–139, 2001.

[30] K. Pawlikowski, J. Jeong, and R. Lee. Letters to the editor. *IEEE Communications Magazine*, pages 132–139, 2002.

[31] K. Pawlikowski and V. Yau. *Akaroa-2 Manual*, 1995. Department of Computer Science, University of Canterbury, ChristChurch New Zealand. URL: http://www.tkn.tu-berlin.de/research/ns-2_akaroa-2. Page accessed on April 21, 2005.

[32] G. Pei, M. Gerla, X. Hong, and C. Chiang. A wireless hierarchical routing protocol with group mobility. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1538–1542, 1999.

[33] L. F. Perrone, Y. Yuan, and D. Nicol. Modeling and simulation best practices for wireless ad hoc networks. In *Proceedings of the 35th Conference on Winter Simulation*, pages 685–693, 2003.

[34] The CMU Monarch Project. The CMU monarch extensions to the ns simulator. URL: http://www.monarch.cs.cmu.edu/. Page accessed on April 25, 2005.

[35] The SEACORN Project. SEACORN simulation tools. URL: http://seacorn.ptinovacao.pt/sim_tools.html. Page accessed on April 23, 2005.

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

[37] N. Sadagopan, F. Bai, B. Kishuamachori, and A. Helmy. PATHS analysis of path duration statistics and their impact on reactive manet routing protocols. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 245–256, 2003.

[38] S. Sanchez. Output modeling: abc's of output analysis. In *Proceedings of the 33rd Conference on Winter Simulation*, pages 30–38, 2001.

[39] R. Sargent. Verification, validation, and accreditation of simulation models. In *Proceedings of the 32nd Conference on Winter Simulation*, pages 50–59, 2000.

[40] L. Schruben. Detecting initialization bias in simulation output. *Operations Research*, pages 569–590, 1982.

[41] R. Zheng, J. Hou, and L. Sha. Asynchronous wakeup for ad hoc networks. In *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pages 35–45, 2003.