Comparing Simulation Tools and Experimental Testbeds for Wireless Mesh Networks

Kefeng Tan, Daniel Wu, An (Jack) Chan, Prasant Mohapatra
Department of Computer Science
University of California, Davis, CA 95616
E-mail: {kftan,danwu,anch,pmohapatra}@ucdavis.edu

***Abstract*—Wireless simulators provide full control to researchers in investigating traffic flow behavior, but do not always reflect real-world scenarios. Although previous work pointed out such shortages are due to the limitation of radio propagation models in the simulators, it is still unclear how these imperfect models affect network behavior and to what degree. In this paper, we quantify network behavioral differences between simulations and real-world testbed experiments. We compare and analyze the experimental results from indoor and outdoor experiments with the results from NS-2 and Qualnet simulations. We find that in the PHY layer, the distribution of received signal strength in experiments is usually different from simulation due to the antenna diversity. However, path loss, which is regarded as a dominating factor in simulator channel modeling, can be configured to match the real-world behavior. For the MAC layer, increasing traffic load on a flow may cause significant performance degradation in experiments, but it is not the case in simulations. Interference is inadequately captured in simulations and cannot show the flow level unfairness phenomenon. In the network layer, a few dominant routes exist in testbed experiments while routes in simulations are less stable. These findings give the wireless research community an improved overview about the differences between simulations and testbed experiments. They will help researchers to choose between simulations and experiments according to their particular research requirements.**

I. Introduction

In the past decade, due to the emphasis on practicality, experimental approaches are more popular in wireless networking research than simulations. Many of the seminal work, such as Roofnet [1], MeshNet [12], ORBIT [17], and QuRiNet [23] are based on testbed experiments. On the other hand, because of hardware limitations and the high cost in testbed experiments, simulators still play an important role in various networking protocols development [4], [5], wireless system quality of service (QoS) performance evaluation [3], and on-line traffic analysis [8].

Simulations have advantages that can never be replaced by testbed experiments. In simulations, network scenarios can be easily constructed and modified, and data can be easily collected. More importantly, simulations can model large scale network topologies that could be very expensive, if not impossible, in testbed experiments. Building a wireless network testbed would require hardware and labor resources. Moreover, testbed experimental results are heavily affected by the testing environment, that is often highly random and uncontrollable. For example, wireless communication between two parties can be affected by objects passing by or by the weather condition.

However, wireless network simulators have their own limitations. Due to the abstracted physical-layer modeling, simulators are often accused of not providing trustworthy real-world results. In this paper, we focus on investigating the discrepancies between simulations and testbed experiments. We aim at finding how the imperfections of channel modeling in simulators affect network behaviors and to what degree. In particular, we study multihop traffic flow situations in testbed experiments and simulations to compare their differences. We also study the routing stability and route distributions in testbed experiments and simulations. We find that dominant routes in simulations are less dominant, and routes are short-lived, compared to those in real-world networks.

Our contributions are three-fold:

- We give a systematic and comprehensive comparison between simulations and experiments. The studies are based on multihop wireless networks.
- In each of the PHY, MAC and routing layers, we find the differences that affect the network performance for simulations and experiments. In the PHY layer, we find that simulators do not provide antenna diversity which is usually utilized in nowadays Wi-Fi devices. [1] However, path loss, which is regarded as a dominating factor in simulator channel modeling, can be configured to fit real-world scenarios. In the MAC layer, if hardware processing speed is slow, a serious flow level unfairness could occur when the traffic load is large and competing flows exist. But this phenomenon is not captured in wireless simulators. In the network layer, a few dominant routes exist in testbed experiments while in simulations routes are less stable.
- We give a clear picture about the major differences between simulators and testbeds, and how these differences affect the network traffic flow behavior. It helps researchers weigh the pros and cons in choosing between simulations and testbed experiments when carrying out wireless network studies.

The rest of paper is organized as follows. In Section II, we

[1] Most modern Wi-Fi devices are equipped with multiple antennas and automatically select the antenna with the strongest received signal strength to use. It is worth noting that antenna diversity here is different from the MIMO (multiple-input multiple-output) technique in IEEE 802.11n [16].

978-1-4244-7265-9/10/$26.00 ©2010 IEEE

briefly introduce the background and related works. We then describe the experimental setup in Section III. The impacts of the differences between simulations and testbeds in the PHY, MAC and routing layers are detailed respectively in Section IV, Section V, and Section VI. We conclude the paper in Section VII.

II. Background and Related Work

NS-2 (network simulator 2) is one of the most popular wireless network simulators. Due to its open source nature, many academic researchers use it. After the IEEE 802.11b PHY/MAC model was introduced into NS-2, many enhancement modules (i.e., IEEE 802.11a/g/e, multirate schemes, energy consumption models) were added. QualNet is a popular alternative to NS-2. It was developed from the GloMoSim simulator [25]. Compared to NS-2, QualNet considers more factors that potentially affect the network performance (i.e., antenna parameters, weather factor) . In addition, QualNet supports parallel simulation and provides easy access to some of the more commonly used statistical metrics.

However, simulators are still plagued by inaccurate modeling of wireless networks. One of the more prominent problems is inadequate modeling of the wireless physical layer [2], [11]. Specifically, assumptions in simulators are not always valid in reality. For example, simulators model signal strength as a simple function of distance. Due to the random nature of the wireless physical channel, there is no perfect physical model which can capture all the factors currently. However, simulators are still very useful as they provide full control to researchers in studying traffic flow behavior. Simulations also reduce the cost, in terms of both time and money, for a lot of networking research. Therefore, studying the detailed difference between simulations and real-world testbed experiments can provide important insights for the networking research community.

Previous works have discussed wireless network behavior in the real world. For example, Ratul et al. [14] analyzed the MAC behaviors and Krishna [15] compared the routing stability in Roofnet and MeshNet, but neither of them compared their results with simulation results. The work in [9] also compare traffic flow behavior in simulations with those from testbed experiments. However, the study is limited to an indoor network and evaluated NS-2 simulator only. More comparisons are expected over NS-2, QualNet, and testbed experiments on different hardware platforms in a wide range of network scenarios. The scenarios should include single flow, multiple flows, indoor, and outdoor settings.

In our paper, we provide a comprehensive and systematic comparison between simulators and experiments. In particular we study the impacts of multihop, transmission rates, interference, as well as the discrepancies in routing stability.

III. Methodology

We take an experiment-driven approach in this work. We designed and deployed testbeds for different network scenarios and replicate them in simulators. We compared metrics which include goodput, delay, and statistical results (i.e., probability distribution of dominant routes) to measure the differences between experiments and simulations. The factors we analyze include path loss and antenna diversity (PHY layer); traffic load, bit rates, number of hops, flow fairness and interference (MAC layer); and route length, persistence, stability and diversity of routes (routing layer). To analyze how a particular factor affects the system performance, we fix all other parameters to minimize their impacts. For example, to find out how different rates affect a multihop flow, we use static routing to avoid potential influences from different routing protocols. We describe the testbeds, simulations, and scenario traffic flows next.

A. Experiment Testbeds

To capture the impacts of different processing capabilities, buffer sizes, and number of antenna on network performance, we use different hardware for the different testbeds.

Soekris: We build a wireless mesh network. We use Soekris boards as the mesh routers. They are 266 MHz x86 Soekris net4826 embedded devices running custom built Linux distribution with 2.6.23 Linux kernel [19]. Each of them has 128MB SDRAM main memory and 64MB compact flash, and equipped with one antenna. Associated clients with the mesh routers are HP nc6000 laptops running Linux kernel of 2.6.25. Each client also has a wireless card with an Atheros chipset for wireless connection to the router. MadWifi (version 0.9.4) is installed in all nodes, including routers and clients. We deliberately configure the wireless mesh network with IEEE 802.11a to reduce interference from existing 802.11b/g networks. All nodes are operating on channel 36 (5.18GHz). This platform represents a series of low-end 802.11 devices with one antenna, and have limited processing capability and buffer size.

The Soekris testbed is used for indoor experiments. As shown in Fig. 1, the testbed is composed of nine nodes. The squares represent the laptops, while the circles are the Soekris boards. There are two competing parallel flows. This experiment is done on the second floor of a three-story office building, where concrete walls create non-line-of-sight transmission environment. The solid line indicates the connectivity among nodes. By *connected*, it means a node can directly (without routing) *ping* another node in the network with more than 50% successful rate at 6 Mbps. Otherwise, the link between them is defined as *unconnected*. The interference range, from our measurement, is roughly two-hops away.

Laptops: In this testbed, we use laptops only to form a mesh network platform. Laptops are HP model nc6000. Each of them is equipped with an Intel Pentium M 1.6 GHz processor with 512 MB DDR SDRAM, and an HP W500 802.11a/b/g wireless LAN card with an Atheros chipset. The operating system is Linux with kernel version 2.6.25 and WLAN driver is MadWifi (version 0.9.4). As in Soekris, all experiments are conducted on Channel 36. This platform represents a set of modern 802.11 devices with dual antennas (one on each side of

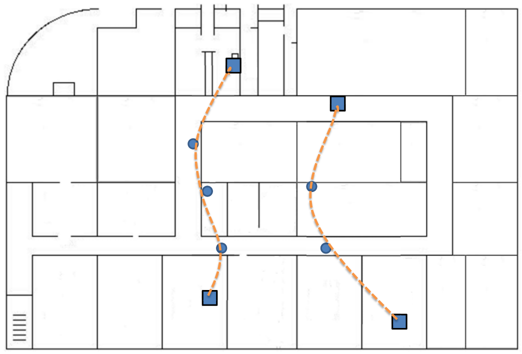


Fig. 1. Soekris indoor mesh network topology.

the LCD screen in the laptop), and have sufficient processing capability and buffer size.

We carry out both indoor and outdoor experiments on this laptop-based mesh network testbed. The outdoor part is performed on an open grass field, so all transmissions are line-of-sight. The indoor part is conducted in a reading room of a public library, where bookshelves are neatly arranged in the center, surrounded by reading desks and chairs. The room has the width of 32 m and has the length of 51 m. In the indoor scenario, transmissions among nodes are non-line-of-sight.

QuRiNet: To collect the network performance results from diverse environments, we also perform some routing experiments in QuRiNet, an outdoor mesh network [23]. QuRiNet is located in the Quail Ridge Natural Reserve, California. It consists of more than 30 nodes and provides wireless coverage over 2,000 acres of hilly terrain. All the access points in QuRiNet are dual radio and utilize multiple channels to achieve high goodput. For details of QuRiNet, interested readers can refer to [23].

B. Simulations

We use NS-2 (version 2.34) and QualNet (version 4.0) as target simulators [7], [20]. They are the two most commonly used wireless network simulators. We run all the simulations on them. In each simulation, we configure the simulation parameters to represent the values observed in the real network testbed. Due to space limitations, these parametric values are not individually presented here.

C. Traffic

Although it is tempting to conduct an overall comparison for both TCP and UDP traffic, we have found that TCP traffic would incur a dramatic throughput variance even during consecutive experiments. This is because that TCP has congestion control and its recovery algorithm changes packet transmission time, and the throughput is heavily affected. For a fair comparison, and avoid this auto feedback control loop in TCP, we use UDP traffic only. For the similar reason, we do not use auto rate adaptation schemes on MAC layer. Iperf (version 2.0.2) is used to generate application level traffic in all experiments [21].

D. Running Time for Simulations and Experiments

For each simulation, each test runs for five minutes, and is repeated for five times. For experiments on testbeds, each test runs for two minutes, and is repeated for five times as well. Ideally, the longer the running time, the more accurate results we have. From the observations, given a fixed environment and a set of parameters, we find that the network performance is quite stable. Therefore, we only run each test in the experiment for two minutes to save time for other tests. Unless we state explicitly, all the results presented in this paper are averaged values of multiple runs.

IV. PHY: Beyond Inaccurate Channel Modeling

Channel model plays a fundamental role in simulators for realizing PHY layer of wireless networks. However, all channel models are mathematical functions and a set of formulas. Radio propagation in reality does not always follow the functions and formulas in simulators. Previous research showed this problem by evaluating the received signal strength (RSS), the probability of symmetric beacon, the reception ratio at different distances [11]. We performed similar experiments and our results support the previous conclusions. However, other factors affect the results as well. Engineering design parameters, such as antenna diversity, also play an important role in the differences between simulations and testbed experiments.

A. Antenna Diversity

It is well known that physical modeling cannot accurately capture all the randomness in real wireless channels. However, from our study, we find that another important factor that has been missing from simulators is antenna diversity. Antenna diversity is a widely adopted technique in modern 802.11 devices. It takes advantage of the fact that quality of received signal at two antennas can differ a lot if they are spaced at least one wavelength apart (12.5 cm at 2.4 GHz and 5.79 cm at 5.18 GHz). Based on this observation, the antenna with the best signal quality will be automatically selected for transmitting or receiving frames. It is worth noting that the antenna diversity here is different from the antenna diversity technique in MIMO (Multiple-Input Multiple-Output). On reception of a frame, MIMO would further utilize advanced signal processing technique to effectively combine received signals from multiple antennas, while the antenna diversity in our devices just simply picks up one of the best. Today's laptops usually have one antenna installed on each side of the LCD screen. Antenna diversity is also supported by MadWifi drivers [13].

Method: We use *wireshark* [22] to record packets from the **Laptops** testbed (in which each node has two antennas). In this experiment, the bitrate does not affect the received SNR. We simply use 6 Mbps bitrate. Then we analyze the trace files of the recorded packets. In each frame's record, the received SNR

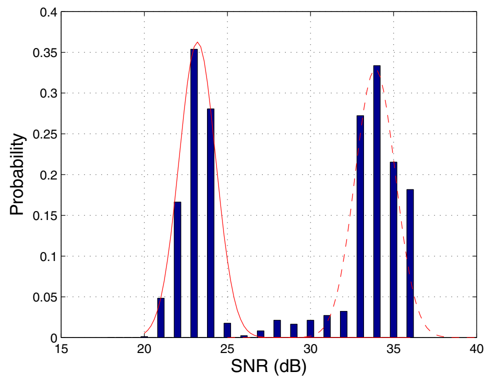


Fig. 2. Example of indoor SNR PDF: the plot shows the distribution of 123,744 frames received by a laptop in 52 minutes.

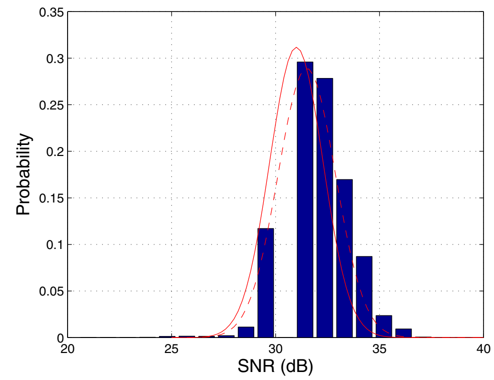


Fig. 3. Example of outdoor SNR PDF: the plot shows the distribution of 135,295 frames received by a laptop in 52 minutes.

(in dB) and the antenna from which it was received have been traced. We then separate the frames by antenna, and plot their PDFs. Different from a previous work [18], in which two wifi interfaces are exploited to trace frames' SNR, laptops in ours are only equipped with one 802.11 interface but two antennas. Other works such as [6] also discuss the antenna diversity, but we conduct both indoor and outdoor experiments and provide a larger volume of data to highlight the difference in SNR brought by antenna diversity. We find the higher SNR antenna received more samples than from the lower SNR antenna due to the transmitter placement and signal propagation effects.

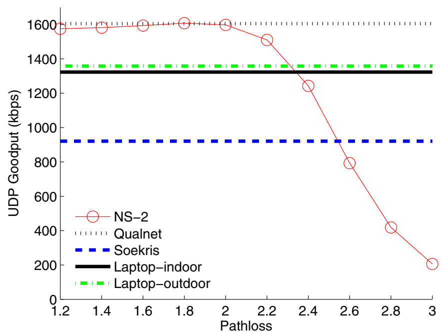
Results: Fig. 2 and Fig. 3 are two examples of a receiver's SNR patterns from a single laptop. Fig. 2 shows a bi-modal Gaussian distribution of received signal's SNR in a multipath indoor environment, where one antenna has higher SNR mean value than the other. However, this phenomenon is less obvious in the outdoor environment, as shown in Fig. 3. The outdoor nodes have a strong line-of-sight transmission that allows both antennas to have a balanced received signal strength. In addition, we find that there are plunges on the verge of Bell-shape in both figures. The antenna selection scheme accounts for this. Once the received signal's SNR falls into the verge, it is highly possible to trigger an antenna switch. Therefore, this low SNR sample is not being traced. On the other hand, received signal's SNRs in simulations are generated based on formulas, whose PDF follows the distribution which the user configured. For instance, if Rayleigh model is used, the received signal's SNRs follows Rayleigh distribution.

B. Configuring the PHY Channel

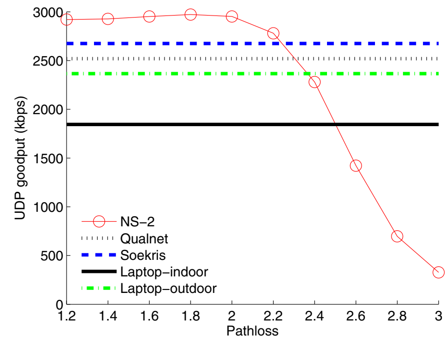
Wireless channel modeling is probably the most important component at the physical layer. Both NS-2 and QualNet have provided some common physical models: freespace, two-ray ground, shadowing model, Rayleigh fading or Ricean fading. Additionally, QualNet supports richer libraries than NS-2: Irregular Terrain Modeling (ITM), High speed fading, etc. For comparison purposes, we use two-ray ground model plus shadowing model in both simulators. Nevertheless, the dominating factor is the path loss, which essentially determines channel gain. Bigger path loss leads to a stronger signal loss. In NS-2, path loss has to be set as a global parameter by users to control channel quality, while there is no such a parameter in QualNet. Instead, in QualNet, an optional interface for an user-defined channel file is provided, and the individual channel between each pair of nodes can be configured. Since our goal is to find out the differences between experiments and the most commonly used scenario in simulators, we set the path loss in NS-2, but keep the default channel settings in Qualnet (path loss is 2.0 and variance is 4.0). Through these tests, we find that path loss, which is regarded as a dominating factor in simulator channel modeling, can actually be configured to fit real-world scenarios.

Method: The network topology of both **Soekris** and **Laptops** (indoor and outdoor) is H-shape - two parallel linear topology. Over each, a single UDP traffic flow is loaded. We also evaluate their performances at different bit rates: 6 Mbps, 12 Mbps, 24 Mbps and 48 Mbps. Settings in simulators are the same, except we change the path loss in NS2 to see how goodput varies.

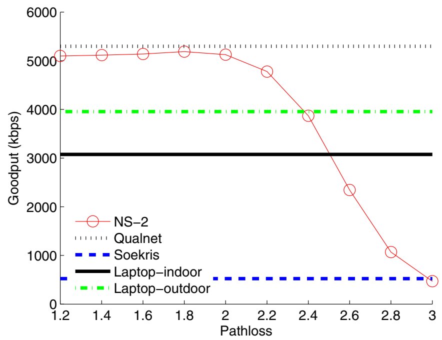
Results: Graphs in Fig. 4 from left to right show the goodput at different bit rates. In each graph, experiments and QualNet result are constant, while NS-2 result changes with the path loss. In our experiments, the performance of the laptop-based outdoor scenario (laptop-outdoor) is better than that of the laptop-based indoor scenario (laptop-indoor) due to line-of-sight transmissions. Soekris and the laptop-indoor are both tested in indoor, but they are in the different environment (the former is in an office building while the latter is in a library). The Soekris' performance does not seem to be consistent, it can be attributed to its multipath fading environment. Especially, at high rates, links between Soekris nodes become extremely vulnerable and therefore the goodput goes down to zero. In simulations, the QualNet results are almost always better than those in the experiments due to its benign channel setting. The NS-2 results match laptop-indoor results most closely when the path loss equals to 2.4 for indoor and about 2.5 for outdoor. Even if we change the transmission bit rate, this setting still matches very well. In the following



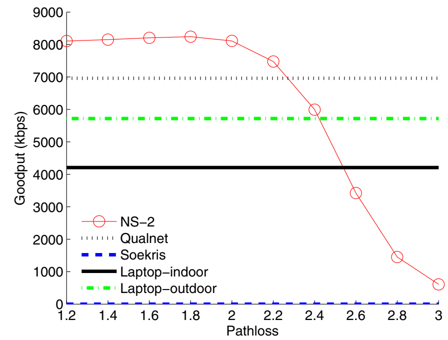
(a) Bit rate: 6 Mbps



(b) Bit rate: 12 Mbps



(c) Bit rate: 24 Mbps



(d) Bit rate: 48 Mbps

Fig. 4. Throughput varies with channel path loss at different bit rates

comparisons, we always use these path loss settings in NS-2 if not specified otherwise.

V. MAC: Flow-level Unfairness Due to Interference

In this section, we first compare network performance at various bit rates. Secondly, we compare the impacts from the intra-flow interference and inter-flow interference. We discover that the inter-flow interference can result in extreme unfairness at the flow level. However, such phenomenon cannot be captured properly in simulators.

A. *Impact of Transmission Data Rate*

In Section IV-B, we have seen the network flow performance at certain transmission bit rates. Fig. 5 presents the flow performance across all eight bit rate schemes in IEEE 802.11a. At low rates, almost all results increase linearly. This increase becomes slow because the link becomes unreliable at high rates. NS-2 results (path loss is set to 2.5 for outdoor environments) matches the laptop-outdoor results closely. However, due to the bigger variance in laptop-indoor results, they do not match even if we reconfigure the path loss parameter to fit the indoor environment. The dramatic plunge at 54 Mbps in laptop indoor is not captured in simulators (NS-2 and QualNet).

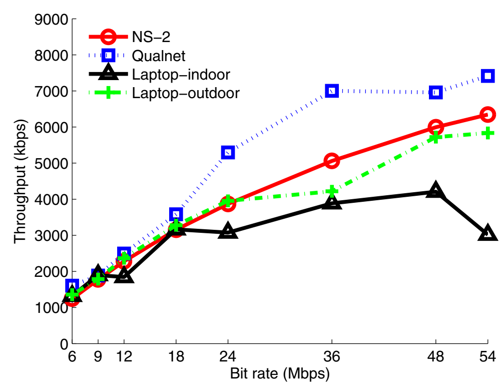


Fig. 5. Throughput varies with bit rates

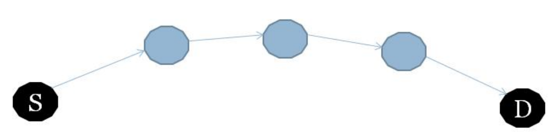


Fig. 6. Linear network topology for a multihop network.

B. *Impact of Multihop*

Now, we investigate the performance differences caused by multiple hops in mesh network.

Method: We first let the sender transmit constant bit rate

(CBR) traffic to the nodes two hops away, then to the nodes further away sequentially. The nodes are closely placed and reside in the same collision domain. In this case, the interference comes from the flow itself, we call it "intra-flow interference". The data and ACK bit rate are set to 6 Mbps to ensure the wireless links remain reliable. The packet size at the application layer is 1460 byte. We also vary traffic load from 1.5 Mbps, 2 Mbps, to 2.5 Mbps. Fig. 6 shows five Soekris nodes forming a linear topology in an office environment.

Results: The results are shown in Fig. 7 through Fig. 12. In each figure, we show how goodput or delay vary with the number of hops. Note the delay is end-to-end delay, which includes MAC layer delay and queueing delay. When the number of hops is 2 or 3, the results in NS-2, QualNet and the testbed match each other closely. However, they diverge as the number of hops or traffic load increases. A conservative channel setting can explain consistently high goodput and low delay in QualNet. But, the reasons accounting for the testbed results' drop are less straight forward. As the analysis in [24], packet loss in a mesh network can be attributed not only to network-induced factors such as interference and wireless channel, but also to node-induced factors like processor queue in the operating system. If the processing speed at each node is not fast enough to handle incoming packets, nodes will drop packets even if they were correctly received. This is not a problem for simulators, but can be a serious issue in testbeds. When traffic load is lower than network capacity, nodes can handle the traffic, but if the traffic or number of hops increase, it results in queue overflow. This is exactly the case in our testbed since each Soekris board have very limited computational resources.

C. Flow Unfairness Under Interference

The above experiments were performed in a single collision domain, where only intra-flow interference occurs. In this section, we investigate how well inter-flow interference is captured in simulators. Strictly speaking, interference modeling is a physical layer problem. Since our experiments involve flow level wireless medium contentions, we put the discussion in the MAC section. The interference modeling problem has been discussed in [10], in which the authors pointed out the problem of resolving collisions in simulators. Neither NS-2 nor QualNet follows the physical layer capture technique found in the real-world. The latest models in both simulators use signal to interference and noise ratio (SINR) to solve physical layer capture issues.

Method: We are interested in interference impacts on flow level performance. To this end, we construct scenarios with two flows transmitting simultaneously. Impacts of inter-flow interference are quantified as follows. The performance of a single flow is first measured, and then the interfered performance is tested. Dividing the latter by the former, we can get the degradation in percentage. For example, 10% goodput degradation means that the interfered goodput is down to 10% of that in non-interfered case.

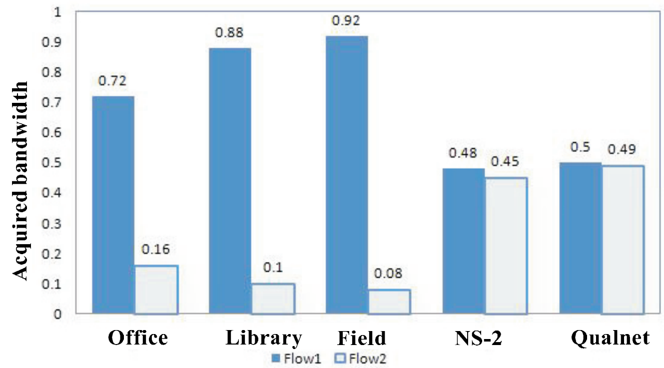


Fig. 13. Flow throughput fairness comparisons

We performed experiments on two platforms, in three environments (office building, open library room, and an outdoor field), and over two network topologies: X-Shape and H-Shape. In both the indoor and outdoor environments, we use eight laptops to form two three-hop flows. In each flow, nodes are placed about 13 meters away to ensure transmission results at high data rates can be measured. The accurate distance for carrier sensing and interference is hard to measure. Especially, effectiveness of interference is related to bitrate as well. Since 13 meters is about the maximum distance to support 54 Mbps between two laptops, we double that distance to create an inter-flow interference. Two parallel flows are set about 25 meters apart. The PHY settings in simulators do not affect the flow-level competition.

Results: An extreme unfairness situation between flows can be observed in Fig. 13, where we average the results from all tests. In contrast, the simulators assume flows are identical if the transmit power, the distance, and the channel quality between nodes are identical.

Analysis: In our testbed experiments, the hardware transceiver's capability determines each flow's performance. The hardware processing speed discussed in section V-B is not a concern here since we do not overload the flows. Instead, the difference in transmitting power, receiving sensitivity, and careeer sensing threshold at each node contributes to the flow level performance differences. Given identical non-interference environments with the same distance between nodes, if one flow consistently has lower goodput than another, that indicates some of nodes in this flow have weaker transceiving capability. When this flow comes into competition with the other flow, the difference is magnified. The stronger flow is slightly affected by the interference from the weaker one. The weaker flow can only deliver a small amount of packets, which introduces little interference.

Topologies and environments are not the dominating factors in interfered scenarios. Previously, we found a performance difference between indoor and outdoor environments in Section V. However, when interference exists, that difference becomes negligible. The H-shape and X-shape topologies do not introduce much difference either.

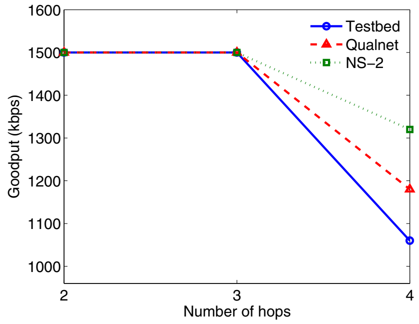


Fig. 7. Traffic load 1500k: goodput

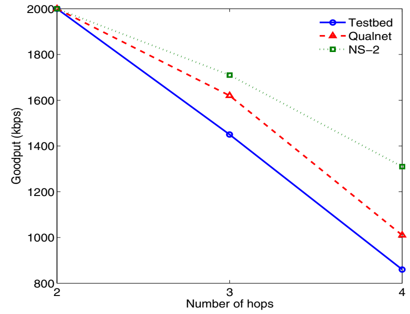


Fig. 8. Traffic load 2000k: goodput

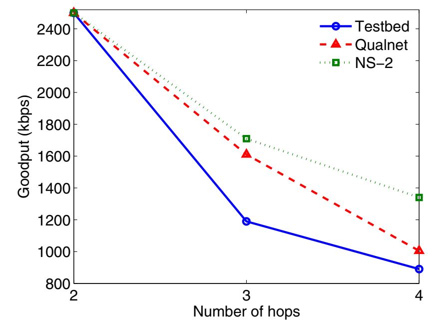


Fig. 9. Traffic load 2500k: goodput

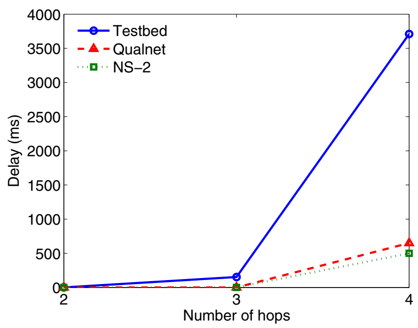


Fig. 10. Traffic load 1500k: delay

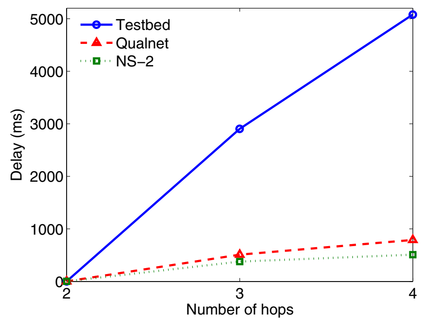


Fig. 11. Traffic load 2000k: delay

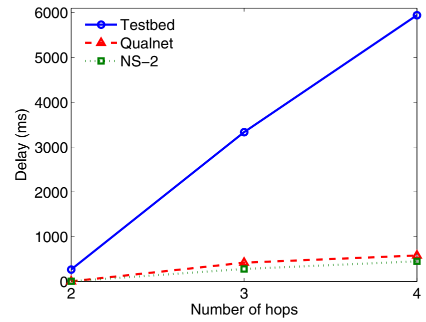


Fig. 12. Traffic load 2500k: delay

VI. Routing: Route Stability

In this section, we quantify the differences in routing stability between simulators and real-world networks. The meaning of this study has two-folds. First, it can provide a better understanding on simulation results of wireless routing protocols. Although most of the existing routing protocols do not consider the routing stability, routing instability, however, can lead to routing pathologies like packet reordering. Second, routing stability affects simulation results of wireless mesh network management, planning, and radio placement. For example, if routes change slowly, some centralized channel or radio assignment might be able to work. Otherwise, slow convergence in centralized scheme makes it impossible.

A. Network Settings

We use OLSR as the routing protocol in both our experiments and simulations. In simulation, we record routing table changes every 0.1 seconds for a given source-destination pair. Since there are a large number of possible source-destination pairs, we pick up 30+ source-destination pairs that are distributed across the whole network. For experiments in QuRiNet, we probe the route between a source-destination pair every 5 seconds to avoid overloading the network. The low recording frequency may miss route updates, but our results will show that they do not change the trend in discrepancy.

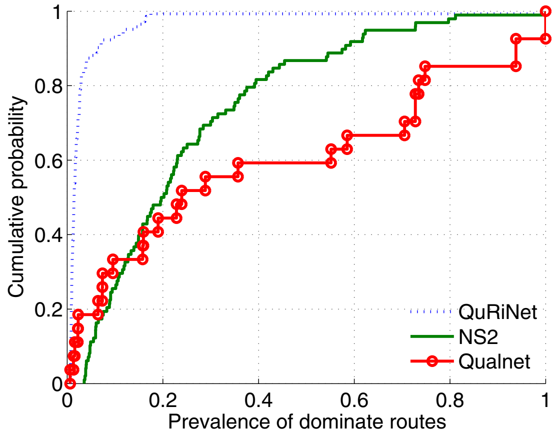
QuRiNet is a heterogeneous network, and nodes are equipped with omnidirectional and/or directional antennas. Nodes are configured to use different transmit powers. Instead of trying to reconstruct the exact same network in the simulators, we replicate the network's connectivity. From an arbitrary node in either network or simulators, it should have similar reachability to all other nodes. We argue this should be enough for our purpose because OLSR use expected transmission times (ETX) as the routing metric and the background traffic in QuRiNet is low enough that it does not affect ETX. Therefore, the dominating factor for routing stability is the channel quality and channel stability. We collect four-days (96 hours) of data from QuRiNet and more than 2 GB data from the simulations.

B. Route Prevalence and Persistence

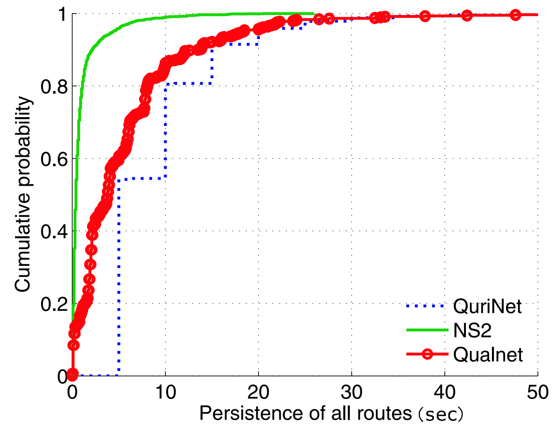
Method: Two metrics are used to analyze routing stabilities, *prevalence* and *persistence*. The *prevalence* is the normalized occurrence frequency: occurrence of a route over that of the *dominant route*. The *persistence* means the duration in seconds for which a route lasts before it changes.

For a given source-destination pair, we analyze its routing prevalence. The *dominant route* is defined as the one being observed most often over all records. If the dominant route occurs n seconds in total, and another route lasts m seconds cumulatively, we define the *prevalence* as a ratio $p = \frac{m}{n}$, and the range is (0, 1]. For example, a route's prevalence is $\frac{1}{50}$ means that the dominant route appears 50 times more often than this route. This relative prevalence can capture the relationship among all routes.

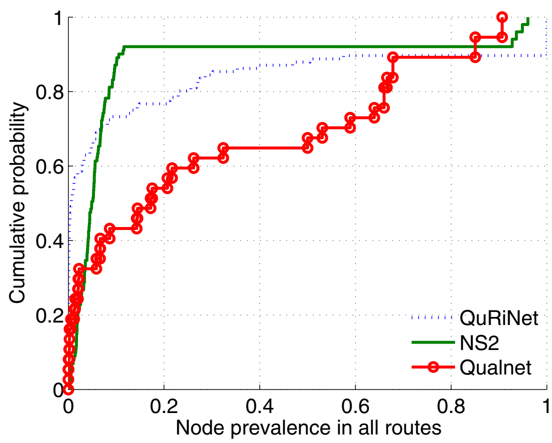
Results: Fig. 14(a) shows the cumulative distribution of the route prevalence over all transmission pairs. For clarity, routes that occur three orders less than the dominant route have been omitted. We observe that the dominant route occurs



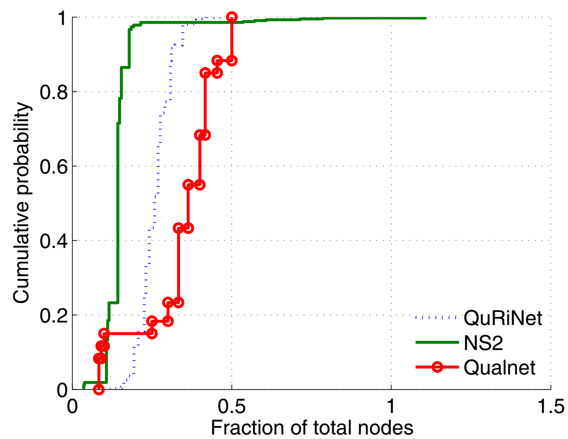
(a) Prevalence of dominant routes



(b) Persistence of all routes



(c) Node prevalence in all routes



(d) Fraction of nodes in each route over total nodes

Fig. 14. Impacts on Routing Stability and Prevalence

much more often than the rest in QuRiNet. Most of the routes are merely occurring 20% as often or less as the most popular one. The dominant route often turns out to be the one with the minimum number of hops. This trend is less obvious in both simulators, where the routes' occurrence frequency are distributed more evenly.

The route persistence analysis is shown in Fig. 14(b). Most of the routes are short lived. Similar to QuRiNet, the median persistence of QualNet is 5 seconds, while NS-2 has 80% of its routes last less than 1 second. The frequent route changes is possible even if OLSR HELLO messages are exchanged every 2 seconds among the nodes because message exchange is not synchronized, and every message exchange may induce a route change. In QuRiNet, since each route is recorded every 5 second, some discreteness in time is observed.

C. Spatial Prevalence and Spatial Distribution

In this section, we will compare the spatial prevalence.

Method: Here the *spatial prevalence* is defined as follows. Every time a node appears in a route, including source and destination, we count its occurrence once. The *spatial prevalence* is the number of cumulative occurrences of one node over the number of all records. The source and the destination will appear in every route between them, and their spatial prevalence is 1. This metric can reveal how often a node appear in all routes. In another words, it shows the routes spatial diversity.

We are also interested in the spatial distribution of routes. We calculate the length of routes in all records, then divide the length by the total number of nodes in the network. We want to use this ratio to measure the distribution of the route length. Note for both simulators and QuRiNet, the total number of nodes in the network is the same, thus the metric is comparable.

Results: Fig. 14(c) shows the cumulative distribution of node spatial prevalence, where the x-axis indicates the fraction of all routes, and the y-axis is the cumulative probability of nodes involved. NS-2 exhibits the most spatial diversity, with approximately 17% of routes covering 90% of nodes; followed by QuRiNet with 20% of routes covering 78% of nodes. This distribution almost matches the power law. NS-2 and QuRiNet are similar in terms of route spatial diversity. The routes in QualNet are spatially concentrated with 60% of routes going through 70% of nodes. In other words, the similarity in QualNet is higher than NS-2 and QuRiNet. This is due to the better channel quality and relative stability of channel model in QualNet simulation.

Fig. 14(d) demonstrates the cumulative distribution results of route length. The medians for NS-2, QuRiNet, and Qualnet are 15%, 25%, and 38% respectively. This indicates that the number of hops in each route in QualNet is generally more than the other two. Together with the results of route spatial prevalence in Fig. 14(a), we have found that even if the length of routes in QualNet is longer, they involve less number of nodes. This strengthens our conclusion about high similarity among routes in QualNet.

VII. CONCLUSION

In this work, we performed a study on the discrepancies between simulators and real-world testbeds. While the discrepancies are broadly expected, this paper provides a detailed and systematic quantification. We have done the simulations on NS-2 and QualNet which are the two of the most commonly used wireless network simulators. We then have compared the results with those from experiments on three different testbeds, single antenna embedded system, laptops with modern Wi-Fi devices, and QuRiNet (an outdoor Wi-Fi network deployed in a natural reserve). The factors we studied include antenna diversity, path loss, multihop, transmission rate, interference and routing stability. We summarize the discrepancies between simulations and testbed as follows. First, in addition of inaccurate channel modeling, simulations also do not model the antenna diversity that is commonly featured in nowadays hardware. Second, a dominant factor in channel modeling is path loss. We find that simulations actually can match experiments in some simple environments like outdoor with line-of-sight transmissions when the path loss parameter is configured properly. This match can be held even if multiple different bit rates are used. Third, the transceiving capability in simulation is just a simple function of distance, while in reality, differences in hardware can result in extreme flow level unfairness in interfered scenarios. Last, we have compared the route stability in the outdoor QuRiNet testbed with that in simulations. We find that routes in simulations are less stable and less persistent than in a real network. In this paper, we give a clear picture about the major differences between simulators and testbeds. It provides an essential reference for researchers to weigh pros and cons in choosing between simulations and testbed experiments in doing wireless network studies.

REFERENCES

[1] John Bicket, Daniel Aguayo, Sanjit Biswas, and Robert Morris, "Architecture and evaluation of an unplanned 802.11b mesh network," in *Proceedings of the 11th ACM International Conference on Mobile Computing and Networking (MobiCom '05)*, Cologne, Germany, August 2005.

[2] Y. Sasson D. Cavin and A. Schiper, "On the accuracy of manet simulators. Technical report," Tech. Rep. Technical report, Distributed Systems Laboratory, Oct 2002.

[3] An Chan and Soung Chang Liew, "Performance of voip over multiple co-located ieee 802.11 wireless lans," *IEEE Transactions on Mobile Computing*, vol. 8, no. 8, pp. 1063–1076, 2009.

[4] Chi Pan Chan, Soung Chang Liew, and An Chan, "Many-to-one throughput capacity of ieee 802.11 multihop wireless networks," *IEEE Transactions on Mobile Computing*, vol. 8, no. 4, pp. 514–527, 2009.

[5] Alix L.H. Chow, Hao Yang, Cathy H. Xia, Minkyong Kim, Zhen Liu, and Hui Lei, "Ems: Encoded multipath streaming for real-time live streaming applications," in *Network Protocols, 2009. ICNP 2009. 17th IEEE International Conference on*, Oct. 2009, pp. 233–243.

[6] D. Giustiniano, G. Bianchi, L. Scalia, and I. Tinnirello, "An explanation for unexpected 802.11 outdoor link-level measurements results," in *Proceedings of INFOCOM Minisymposium*, 2008.

[7] VINT group, "UCB/LBNL/VINT Network Simulator – ns (Version 2)," *http://mash.cs.berkeley.edu/ns*.

[8] Fang Hao, Murali Kodialam, and T. V. Lakshman, "On-line detection of real time multimedia traffic," in *Network Protocols, 2009. ICNP 2009. 17th IEEE International Conference on*, Oct. 2009, pp. 223–232.

[9] S. Ivanov, A. Herms, and G. Lukas, "Experimental validation of the ns-2 wireless model using simulation, emulation, and real network," in *Proceedings of the 4th Workshop on Mobile Ad-Hoc Networks*, 2007.

[10] Andrzej Kochut, Arunchandar Vasan, A. Udaya Shankar, and Ashok Agrawala, "Sniffing out the correct physical layer capture model in 802.11b," in *ICNP '04: Proceedings of the 12th IEEE International Conference on Network Protocols*, 2004.

[11] D. Kotz, C. Newport, R. Gray, J. Liu, Y. Yuan, and C. Elliott, "Experimental evaluation of wireless simulation assumptions.," in *In Proc. ACM/IEEE Intl Symp on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2004.

[12] Henrik Lundgren, Krishna Ramachandran, Elizabeth Belding-Royer, Kevin Almeroth, Michael Benny, Andrew Hewatt, Alexander Touma, and Amit Jardosh, "Experiences from the Design, Deployment, and Usage of the UCSB MeshNet Testbed.," *IEEE Wireless Communications Magazine*, 2006.

[13] "Madwifi userdocs: Antenna diversity [technical document]," http://madwifi.org/wiki/UserDocs/AntennaDiversity.

[14] Ratul Mahajan, Maya Rodrig, David Wetherall, and John Zahorjan, "Analyzing the mac-level behavior of wireless networks in the wild," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 75–86, 2006.

[15] Vivek Mhatre, Henrik Lundgren, Francois Baccelli, and Christophe Diot, "Joint mac-aware routing and load balancing in mesh networks," in *CoNEXT '07: Proceedings of the 2007 ACM CoNEXT conference*, New York, NY, USA, 2007, pp. 1–12, ACM.

[16] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A high-performance MIMO OFDM wireless LAN," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 101–109, Feb 2005.

[17] D. Raychaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. . Singh, "Overview of the orbit radio grid testbed for evaluation of next-generation wireless network protocols.," in *Proc. of IEEE Wireless Communications and Networking Conference*, 2005.

[18] Yong Sheng, Keren Tan, Guanling Chen, David Kotz, and Andrew Campbell, "Detecting 802.11 MAC Spoofing Using Received Signal Strength.," in *Proc. IEEE INFOCOM*, 2008.

[19] "Soekris engineering," http://www.soekris.com/.

[20] Scalable Network Technologies, "QualNet," *http://www.scalable-networks.com/products/qualnet/*.

[21] Ajay Tirumala, Les Cottrell, and Tom Dunigan, "Measuring end-to-end bandwidth with iperf using web100," in *Web100, Proc. of Passive and Active Measurement Workshop*, 2003.

[22] Wireshark, "Wireshark," *http://www.wireshark.org/*.

[23] Daniel Wu, Dhruv Gupta, Stephanie Liese, and Prasant Mohapatra, "Qurinet: quail ridge natural reserve wireless mesh network," in *WiNTECH '06: Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, New York, NY, USA, 2006, pp. 109–110, ACM.

[24] Hua Yu, Daniel Wu, and Prasant Mohapatra, "Experimental anatomy of packet losses in wireless mesh networks," in *Proceedings of IEEE SECON*, 2009.

[25] Xiang Zeng, Rajive Bagrodia, and Mario Gerla, "Glomosim: a library for parallel simulation of large-scale wireless networks," *SIGSIM Simul. Dig.*, vol. 28, no. 1, pp. 154–161, 1998.