UCLA
Papers

Title

In Vivo Characterization of a Wide area 802.11b Wireless Seismic Array

Permalink

https://escholarship.org/uc/item/60j0j3xp

Authors

Lukac, Martin
Naik, Vineyak
Stubailo, Igor
et al.

Publication Date

2007-05-05

eScholarship.org Powered by the California Digital Library
University of California

In Vivo Characterization of a Wide area 802.11b Wireless Seismic Array

Martin Lukac, Vinayak Naik, Igor Stubailo, Allen Husker, Deborah Estrin

Abstract

Wireless sensor networks using 802.11b long-distance links enable domain scientists to measure physical phenomena in real time and at a scientifically meaningful scale. Although deployments in this setting are increasing, a characterization of the network-throughput is lacking which makes it difficult to predict the performance of sensing applications on these types of networks. We present a study of the MASE seismic data collection network which includes 64 nodes spread across mountains, valleys, rural plains, and urban environments. The network spans 250km with links as long as 43km and has been operating continuously for more than 2 years. The contributions of this paper are three fold: (a) we present a throughput characterization of wireless links for a network in operational use, (b) we suggest settings for the wireless parameters to improve the performance of the network, and (c) we document the uptime of the nodes and the data yield of the network. Our findings are beneficial for the design of future wireless sensor network deployments using 802.11b.

In our networks widely varied settings, we find that while signal-to-noise ratio is not correlated with throughput or distance, throughput is roughly correlated with distance and does not change significantly over time. At the link-layer, enabling the RTS-CTS handshake decreases the throughput while enabling retries improves the throughput. In addition, the default threshold-based rate adaptation algorithm for the transmission bit-rate is suboptimal. We find that the majority of fault cases are due to system issues and not network issues. The data yield over the course of an eight month period was 78%.

1 Introduction

Wireless sensor networks using 802.11b long-distance links enable domain scientists to measure physical phenomena in real time and at a scientifically meaningful scale. The observed phenomenon dictates the data-rate requirements of the network. In the case of seismic sensing applications the sensing rate is on the order of 10's to 100's of Hz on multiple channels, so the required network data-rate is high. Our seismic application utilizes a dense deployment of broadband seismometers covering hundreds of kilometers to study the structure of the Earth's mantle. The system is dense relative to typical seismic deployments and dense relative to the spatial varriability of the earth's structure. Whereas most seismic stations are separated by 10's or 100's of km, our stations are separted by 5 to 10 km. The large amount of data generated and the distance between stations are well matched with a high bandwidth radio such as 802.11b as opposed to the shorter range and lower power CC1000 or CC2240 radios. The physical span of the network may cover mountains, valleys, forests, etc. where Internet connectivity cannot be guaranteed, so, the seismic data has to be retrieved over a multi-hop mesh network This style of large-scale and large-distance 802.11b network has applications not only in seismic studies but also in connecting sites in environmental observatories such as NEON [2, 12] and WATERS [3] as well as connecting structural health monitoring sensor networks between tall buildings [11].

We have successfully developed, deployed, and used this type of seismic sensor network called the Middle American Subduction Experiment (MASE) [1, 7] broadband seismic network. It consists of over 50 wirelessly connected nodes covering approximately 250km. This network has been in use for over 2 years and will be dismantled in Spring 2007 and moved to southern Peru. To our knowledge this is the first of this type of operational network in scale, granularity, and lifetime. We were able to perform a wireless network study alongside the existing seismic data collection and system management applications making use of the network. This was a relativly unique opportunity to study a new type of system (*in vivo*) under operational conditions during its full lifetime. We use throughput as a metric instead of packet delivery rate (PDR) because it enables us to measure the network characteristics with minimal changes to the network settings and the wireless driver. The use of PDR would require significant changes which would be disruptive to the seismic data collection application: logging of MAC level frames at the transmitter to count the number of transmissions, disabling retries at the MAC layer to avoid retransmissions, use of UDP instead of TCP to avoid retransmissions, and disabling of MAC level ACKs.

Although the 802.11b link characteristics are well studied at various OSI layers, most of these studies have focused on the individual link behaviors measuring the performance of a single link in isolation. Such studies have improved our knowledge of the radio behavior in situations where there is no interference from other nodes. But the link performance is greatly affected in the presence of interlink interference [6]. Given the nature of data collection in high-data-rate seismic networks, interlink interference is the standard. The results of our link behavior study should be relevant to other 802.11b mesh networks.

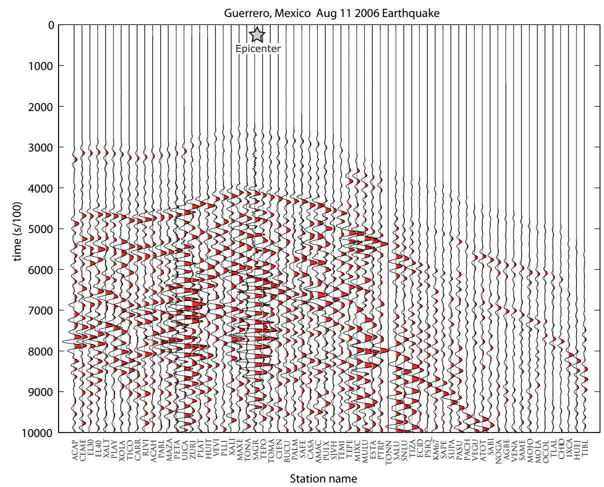

Figure 1: The August 11th 2006 earthquake in Guerrero Mexico. Each point on the X-axis represents the site names for the seismic stations. The Y-axis is time. The first site to register the earth quake was SAGR. If the ground was perfectly homogeneous there would be a even parabolic pattern spreading from SAGR down through the other sites.

Summary of contributions and results:

- The profile of a 64-node wireless sensor network over the central operational period of 8 months is:
  - the average number of working nodes per day is 42
  - the average amount of data received per day is 1.93GBytes
  - the average data yield of network per day 78%
  - in 62% of the faults, the nodes are connected to the network and hence remotely recoverable, and the remaining faults render the nodes disconnected

  The details are discussed in Section 3.

- The *in vivo* characterization of the wireless links is:
  - the median throughput among all the links is about 1Mbps
  - the standard deviation in throughput is equal to or greater than 50% of the mean throughput for half the links
  - throughput is roughly correlated with distance
  - SNR is not correlated with distance nor throughput.

  The details are discussed in Section 4.

- The effects of MAC layer settings for tunning the throughput:
  - the use of MAC-layer retries significantly improve throughput
  - the use of RTS-CTS decreases throughput for low connectivity degree topologies such as ours
  - the default threshold-based rate adaptation algorithm for the transmission bit-rate is suboptimal.

  Although, these MAC-layer behaviors are not inconsistent with theory, we verify them in diverse real environments at a large scale and over an extended period of time. The details are discussed in Section 5.

2 Related Work

Previous research has studied wireless link characteristics for 802.11b radios. The wireless link characterization studies by Aguato et. al. [4] and Chebrolu [6] et. al. are close to ours in terms of network setup, scale. Therefore, we will focus our attention to them.

The dominant experimental methodology in this research was to study link behavior one transmitter at a time. In this methodology, only a single node is allowed to transmit packets while other nodes listen and measure data such as SNR and packet delivery rate (PDR) [4]. The results derived from these classes of experiments improve our understanding of the radio behavior at the physical and link layers. Hence, the results can be use to improve the state-of-the-art of MAC and routing protocols. However given that the inter-node interference has noticeable impact on the link characteristics and performance, the results from [4] do not directly predict link characteristics when there is inter-node interference.

Like the wireless portion of the MASE network, IIT-Kanpur's DGP deployment [6] also uses 802.11b radio, directional antennas, and long-distance antennas. The DGP experiments use a similar methodology to Roofnet but do additional experiments to study the effect of inter-link interference. They studied the degree of interference between two radio channels at the physical layer by measuring the received signal strength when two radios operate in different channels. We complement their work by studying the impact on link characteristics when multiple nodes operate in the same channel, which is relevant to planned deployments of long distance networks.

DTN protocols [8] have been developed for a range of networks that show regular disruptions in links, which precludes the assumption of continuous end-to-end connectivity. The MASE network has frequent and unpredictable disruptions, so we use DTN techniques in the data delivery and system management software. Disruption Tolerant Shell (DTS) [9] is a reliable asynchronous remote shell interface we used to manage the wireless portion of the MASE network. It facilitates reliable dissemination of shell commands to all nodes despite the erratic link qualities and intermittent node disconnection. Our link behavior study analyzes the parameters that affect the link quality which can be used to improve the performance of DTN systems such as DTS.

The application goal of our network is to collect seismic data, so, we will also compare our system with previous seismic wireless sensor networks. The volcano network [13] was a deployment of 16 wireless sensor nodes to monitor a volcano's activity for 19 days. This deployment was targeted to detect major seismic activity in the direct vicinity of the nodes. Their nodes were placed 200-400m apart and they used Tmote Sky nodes, which are equipped with the CC2420 low-power radio with maximum throughput of 250kbps. Using narrow-band sensors, the nodes would detect "interesting" events and report back 60 seconds worth of data. The wireless portion of the MASE deployment consists of 64 nodes deployed over 2 years. The deployment is targeted to record seismic activities at a fine granularity over a long period to understand the structure of the subducted slab along the Cocos Plate as a result of the interaction with the North America Plate. We used XScale based Stargate nodes, which are equipped with 802.11b radios with maximum throughput of 11Mbps. The typical distance between sensor locations is 5-10km, but the distance between the radio links can vary from 100m to 30km. 50 nodes are equipped with broadband sensors and report back all the data. The remaining nodes are relay stations. The MASE network stations used large 12V gell cell batteries while the volcano network nodes ran on D cell batteries.

Figure 1 shows the seismograph of the August 11th 2006 earthquake in Guerrero Mexico. The X-axis represents the IDs of the seismic stations and the Y-axis represents time. The site that first registered the earthquake was SAGR. The seismograph shows the the ground is heterogeneous beacuse if the ground was perfectly homogeneous there would be a even parabolic pattern spreading from SAGR down through the other sites. This pattern would not be identifiable without the fine granularity and high data yield of the MASE deployment.

3 A Year in the Life of the MASE Network

The deployment of the MASE network began mid-year 2005 and it has been operational since then. By March 2006, all the planned nodes had been deployed: 50 total stations with seismometers and addition 14 relay stations without sensors. Over the course of the entire deployment, the total number of operational nodes varied a bit. We profiled the network in terms of the architecture of an individual node, network topology, lifetime, robustness to faults, and data yield.

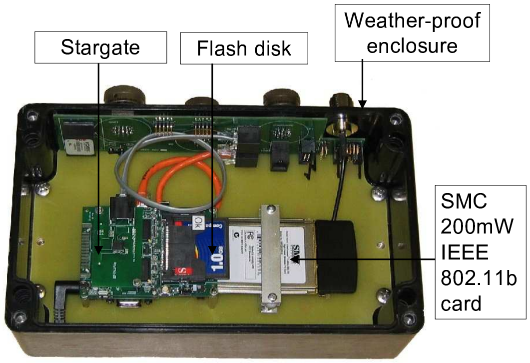

Figure 2: A CDCC: the wireless node which makes up the MASE seismic network. The lid is removed to show the inside components. The wireless card's black antenna is removed and replaced by an external antenna. The four gray circles at the top are the external connectors for the antenna, Q330 (seismic data retriever), ethernet, and serial.

Figure 3: MASE deployment map showing the approximate path of the network. Green dots approximate the path of wireless stations and red dots approximate the standard non wireless stations.

3.1 Anatomy of the Wireless Node

CENS designed and developed a wireless node called CENS Data Communication Controller (CDCC) suitable for long-term deployments in outdoor environments. The CDCC is capable of surviving in the elements without significant down-time. This is important beacuse manually repairing the nodes spread across 250km in Mexico is cumbersome Figure 2 shows a CDCC with its top cover removed. The single board computer used in the CDCC is the Stargate which has 64MBytes of RAM and 32MBytes of internal flash memory. We use 1-4GB Compact Flash cards for storage. The CDCC is equipped with a 200mW high power IEEE 802.11b PCMCIA SMC card with the Prism 2 chipset. The CDCC and the sensing equipment at each station is powered using a 12V cell battery with a capacity of 60 or 70 AmpH charged by a 70W solar panel. The CDCCs are connected to one or two (using a splitter) 2.4GHz directional antennas. Most of the antennas are either 15 dBi yagi (Figure 4) or 24 dBi parabolic (Figure 5) antennas. In some situations 1 or 2 Watt amplifiers are used.

3.2 Large scale, Long Distances, and Diverse Environments

The MASE deployment consists of 112 seismic stations along a 500km line which transects Mexico. 64 of the stations in the deployment are equipped with the wirelessly enabled CDCCs. These 64 wireless stations are deployed along a 250km long line from Cuernavaca through Mexico City stretching towards Tampico. Out of the 64 nodes, 50 nodes are divided into 4 major sections: Huejutla, Cuernavaca, Universidad Nacional Autónoma de México (UNAM), and Pachuca. The nodes in each section form an mesh wireless network, so we have 4 mesh wireless networks. Each of the four sections have one CDCC which is connected to a server via wired ethernet and the server is connected directly to the Internet. Of the 50 stations, a small number are standalone stations which do not have any wireless neighbors which requires the data to be pick up manually once a month.

The nodes are deployed in diverse environments, which range from urban to rural areas. The rural areas traverse mountains, valleys, plains, and fields with dense vegetation. The site locations where chosen primarily based on the seismology requirements: the sites can not be near highways or major roads and the sites must be placed in safe locations. To our knowledge, the MASE deployment is unique in terms of the scale, distance, and diversity of the environment as far outdoor 802.11b networks are concerned. For our wireless experiments we focus on the UNAM, Pachuca, and Cuernavaca sections.

Figures 6, 7, and 8 show the preferred links the seismic data is transferred over to reach the sink. Note that the wireless link lines between the stations do not

Figure 4: A single yagi antennae with solar panel and equipment vault at the HUEJ site.

Figure 5: A parabolic and yagi antenna on top of a tower at UNAM. A splitter is connected to the wireless card to attach two antennas to the same antennae

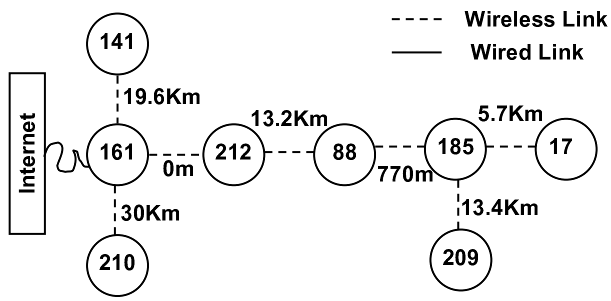

Figure 6: Network topology of the UNAM section of the MASE network.

indicate the only available links between nodes: for instance in Pachuca, node 89 has connectivity with node 100, but the connection is better between 89 and 104.

3.3 Long Lifetime

We measured the lifetime of the network in terms of individual uptime of the nodes. One end to end way to detect that the node is up is if it collects seismic data on any given day. We define an individual node to be up in a day if the node has sampled any data during that day. In Figure 9, we plot the number of up nodes per day, starting from March to October in the year 2006.

We can compare our node uptime to the individual node uptime of the volcano network [13]: while 50% of the nodes were up only for the 50% of the total deployment time for the volcano study, 50% of the CDCCs are up for the entire period of the deployment. In fact, 75% of the CDCCs are up for more than 50% of the deployment time. One main factor contributing to our better uptime is that our nodes run Linux which provides a much more robust software environment with better fault isolation and visibility. Our applications do crash but the crash will not make the node inaccessible, so we can connect to the nodes and diagnose problems remotely. The tradeoff for this more mature, always-on system is that the software is much more complex and does not have the low power deployability of motes.

3.4 Robustness

To determine the reason for data loss in the system we collected the data about the faults observed in the network. We divided the faults into four major conditions depending upon the symptoms of the faults: node is down, node is up but disconnected from the network, node is up and connected to the network, and unknown. A node is deemed to be connected if there is an active route from the base station to the node. If we are unable categorize the faults due to lack of logging, we categorize it as unknown. In Table 1, we list the four major categories of faults and

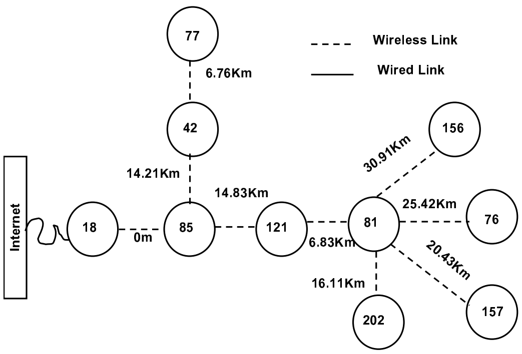

Figure 7: Network topology of the Cuernavaca section of the MASE network.

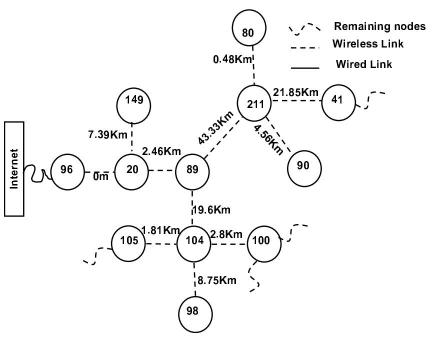

Figure 8: Network topology of the Pachuca section of the MASE network.

give a few examples.

We collect the faults observed from March 2006 to October 2006, a duration of the period is 8 months. Each of the faults is tagged using one of the symptoms mentioned above. The percentage column in Table 1 shows distribution of the 174 observed faults. To determine the number of faults, the archives from the mailing list used to coordinate efforts, to discuss the deployment, coordinate repairing faults, and reporting back on station status was analyzed. The data obtained from the analysis is not complete, but the relative ratios of the types of faults is consistent with what the maintenance teams observed.

The promising characteristics of the faults is that the majority of the faults are curable remotely, i.e. without physically visiting the nodes. Specifically, in 62% of the total faults, the faulty nodes were connected to the network and hence remotely repairable. The remaining 38% of the faults caused the nodes to be unreachable over the wireless network.

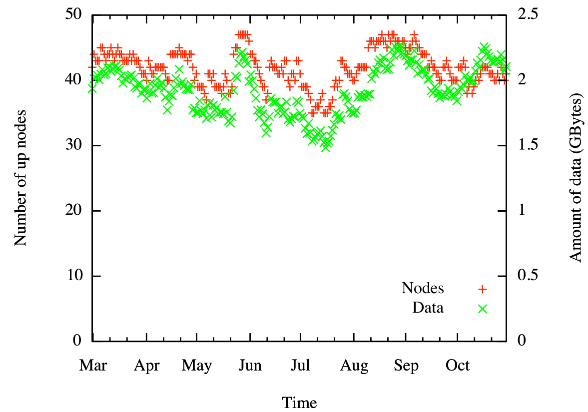

Figure 9: Number of up nodes and amount of data collected per day during eight months in 2006. On average 1.93GBytes of data are collected per day.

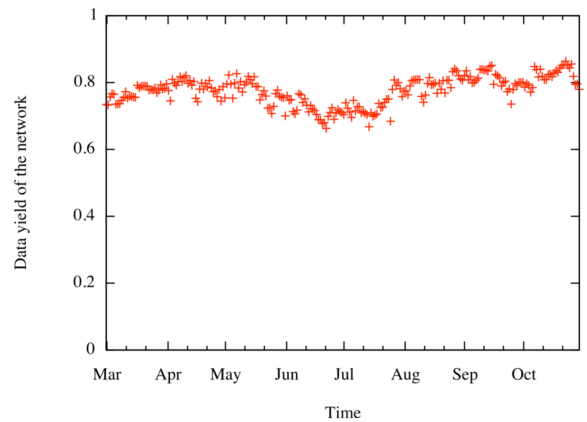

Figure 10: Data yield of the network per day during eight months in 2006. The average yield is 78%.

3.5 Bulk Data Delivery

Each station that is equipped with a seismometer collects data for an hour, bundles and compresses the data into a single file, and sends the data bundle hop by hop in a DTN like fashion to the internet connected server in its section of the network. Each of the seismic data files is 1-3MB in size. In Figure 9, we plot the total amount of data collected from all the CDCC equipped stations per day starting in March 2006 and ending in October 2006. We started in March since not all the nodes were deployed until March. We can see from the plot that, the amount of collected data is directly proportional to the number of the working nodes. The average data collected per day are 1.93GBytes.

Figure 10 shows the data yield of the network. The

Condition	Examples	%
Node is down	corroded connectors, charge controller malfunctioning	10%
Node is up but disconnected from the network	problem with the antenna, burned out amplifier	23%
Node is up and connected to the network	memory card full, configuration problem, file mover crash	62%
Unknown		5%

Table 1: The four fault categories with examples of each and percentages of the total faults.

data yield of the network is defined as the ratio of the amount of data collected at the base stations to the expected total data generated at all the nodes. The data yield is calculated independent of size: each node will generate 24 files, so for a total of 50 sites which generate data, we expect a total of 1200 files per day. A data yield of 1 means that all the generated data is received at the base stations. The average data yield of the network is 78% and remains almost constant. Note that a high volume of data does not necessarily mean a high data yield. For example, although the amount of data collected at the end of May is high as shown in Figure 10, the data yield during the same period is not high. A reason could be that the number of working nodes is high but the network quality is poor.

4 Wireless Characteristics *In Vivo*

There are many factors which can affect the performance of our wireless network, such as distance between stations, types of antennae, number of stations within wireless range of other stations, antenna cable length, use of amplifiers/splitters, and terrain/obstructions between stations. In this paper, we focus on distance between stations. We want to characterize the network under the real conditions and under the real application making use of the network: the delivery of seismic data. This characterization provides an understanding of how well the application can work. Our characterization is performed *in vivo*; we rely on the data already being collected and transferred over the network by the application to provide the characterization. The movement of the data through the network provides us with the TCP

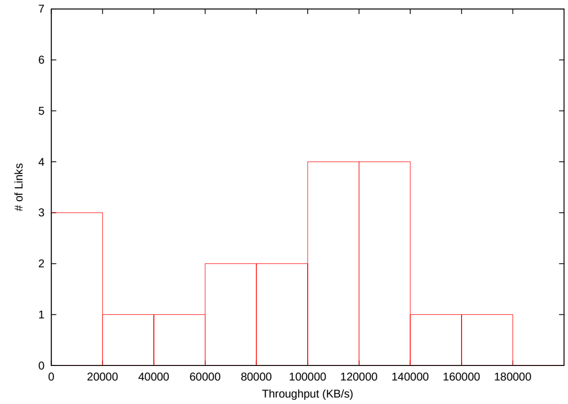

Figure 11: Average throughput for 20 links in the MASE network. The majority of the links have throughput over 100kBps, which is about 1mbps.

link throughput (henceforth referred to as throughput) as well as the link SNR. We neither isolate individual links nor manipulate the data traffic over any link while recording the throughput and SNR measurements. Even without isolating links, our measurements still enable us to fairly evaluate the performance of the links because of the regularity of the data traffic. The nature of the data is that it is the same everyday: each station generates data once an hour and the size of the data stays consistent for each site. Knowing what the network load is lets us compare the throughput of a link across multiple days and across multiple links.

4.1 Methodology

Section 3 described that the hour long 1-3 MB seismic data files are transferred hop by hop using TCP. Each node will only attempt to transfer one file at a time to the next hop node. The throughput is determined by looking at the individual TCP flows of the data that is being transferred through the network and calculating the total number of bytes sent over the duration of a flow per second. We use tcpdump to record the headers of all the TCP flows at each node. The throughput value is measured using the tcptrace program. In addition to the throughput, we record the SNR for each of the the packets using the modified hostap driver from [6]. The resulting logs from tcpdump and the per packet SNR are compressed hourly and transferred to the sink by the same application, which transfers the seismic data. The logs range in size from 10's of kilobytes to a few

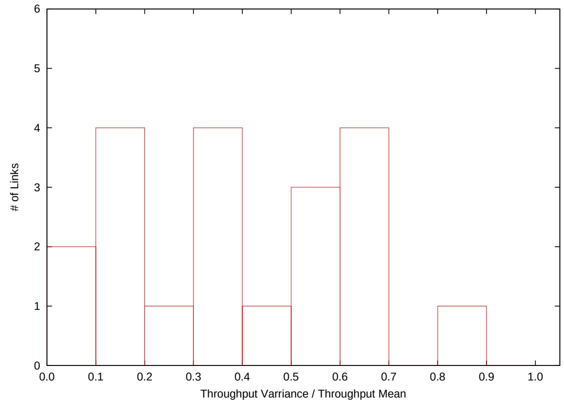

Figure 12: Standard deviation over the mean for 20 links in the MASE network. Half of the links have a standard deviation greater than 50% of the mean.

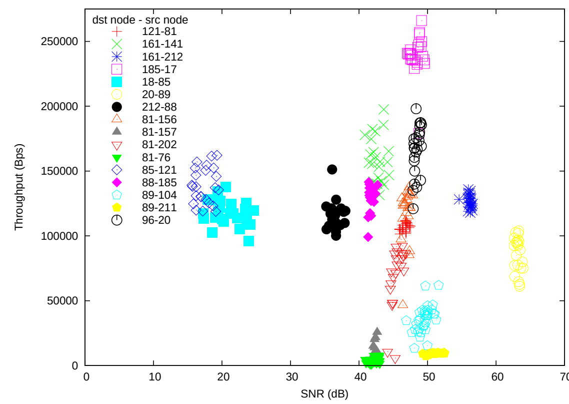

Figure 13: SNR vs. Throughput for 16 links. There are multiple throughputs for SNR between 40 and 50 dB. Similarly, there are many SNR values for throughput between 100kBps and 150kBps.

megabytes and transferring the logs themselves will generate more data in the tcpdump lots and the SNR logs. The nature of the log data is similar to that of the seismic data so we are effectively loading the network with data of similar characteristics.

We chose 20 links for this study. These links were chosen because the majority of the time they had the capacity to carry the seismic data as well as the tcpdump and SNR logs. Other links in the network were poor and unpredictable and we did not want tax them with the extra load from the logs.

We use an HP438A RF power meter to calibrate the received signal strength (RSS) for SMC2532W-B card, as reported by the hostap driver for a received frame at the MAC layer. First, the RF power meter is used to calibrate the transmission power in terms of dBm. Then, we physically connect the transmission port of one wireless card to the receiver port of another card using a cable, whose attenuation is negligible. At the receiving end, we measure the RSS values as reported by the driver per frame. We vary the transmission power and measure corresponding RSS values. Using the previously calibrated transmission power value, we calibrated the RSS values in terms of dBm. According the specification of the hostap driver the RSS values have a linear relationship with the corresponding dBm values. Hence, we interpolate the range of RSS values to dBm values by fitting a line in between them. In order to calculate the SNR for a received frame, we compute a ratio of the RSS and silence values, as reported by the driver, after converting them to dBm values using our calibration.

4.2 Throughput and SNR Over Time

We measure the throughput and SNR over time for the 20 links. For the duration of this collected data, RTS/CTS is off, the number of retries is set to 8, and the bit-rate is set to auto. The total duration of our experiments varies per link: for the majority of the links, it is between 7-14 days and for the remaining links it is either 2-3 days or up to a month. Our overall experience of the throughput is that it stays relatively consistent except for the case where there is some change in the environment that causes a shift in the stable state. For each link we select a relatively consistent 24 hour period and compute the average throughput for all the transfers whose size exceeds 100kB. As we discuss below, the throughput stays stable across multiple days but occasionally changes.

Figure 11 is a histogram of the average throughput for 20 links. Half of the links have a throughput greater than 100kBps, which is a little under 1 Mbit per second.

Figure 12 is a histogram of the ratio of the standard deviation to the throughput for the same 20 links. This ratio allows us to compare the standard deviation across links with drastically different throughput in a fair manner. For example, link A with a average throughput of 100kBps and link B with a average throughput of 20kBps may have the same standard deviation of 10kBps per second. In this situation, link B can be considered as highly varying but A is not necessarily highly varying. In Figure 12, we can see that 40% of the links have a standard deviation greater than 50% of the average link through-

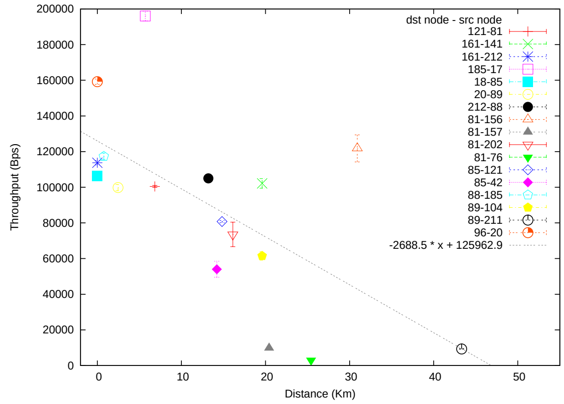

Figure 14: Throughput vs. Distance. We roughly fit a line between throughput and distance.

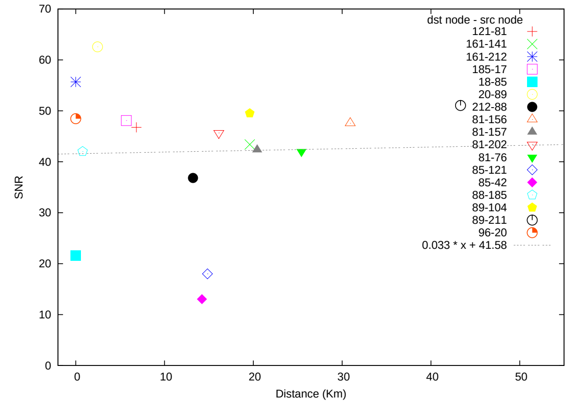

Figure 15: SNR vs. Distance. We do not see a correlation between SNR and distance. For SNR values between 40 and 50, there are multiple corresponding distances.

put. The main reason for this is that the throughput is constantly fluctuating while the link is in a stable state (the average stays relatively constant). This fluctuation is most likely due to the unstable link quality. Over the course of the experiments, a number of the links experience a change in the stable state: the average makes a sudden change, but after that, the average is relatively constant. These jumps most likely occur due to physical or environmental (including surrounding wireless) changes to the station surrounding.

Across the 20 links, the SNR varied greatly: some links are almost as low as 10 dB while others were over 60 dB. The majority of the links have an SNR around 40 dB. The variance in the SNR for all the links is less than 10 dB for 95% of the recorded data. Unlike the throughput, the SNR does not experience any large changes in its stable state.

Figure 13 shows the SNR plotted against the throughput for 16 sites. We used 24-48 hour periods where the average throughput stayed relatively constant. Both the throughput and SNR are averaged by the hour. There is no correlation between SNR and throughput. For many similar values of SNR, there are a wide range of throughputs and for similar throughputs there are a wide range of SNR values. This means that SNR does not play a significant factor in the performance of any individual link.

4.3 Relation of SNR and Throughput with Distance

Using the same data from the previous section, we can compare the throughput and SNR values to the distance. Figure 14 shows the average throughput compared to the distance for each link. There is a rough correlation between the throughput and the distance; the throughput decreases as the distance increases. The r-squared value is 0.4117. Figure 15 shows the average SNR compared to the distance for each link. There is no correlation between the SNR and the distance. This means that SNR is not the primary factor to consider when deploying this type of network.

5 Wireless Tuning

Common tuning parameters at the MAC layer are retries, RTS-CTS, and bit-rate. Depending upon the network settings, the parameters can be tuned for better performance. Usually, retries are beneficial whenever packet losses are high, RTS-CTS reduces hidden terminal effect, and bit-rate selection uses the appropriate modulation techniques to deal with the noisy environment. However, incorrect setting of the parameters can result in unnecessary overhead, thereby degrading the performance. Although theoretical wireless models guide the correct values for the parameters, given the complex nature of the wireless medium, it is not always straight forward to select the correct values. We empirically measure the impact of

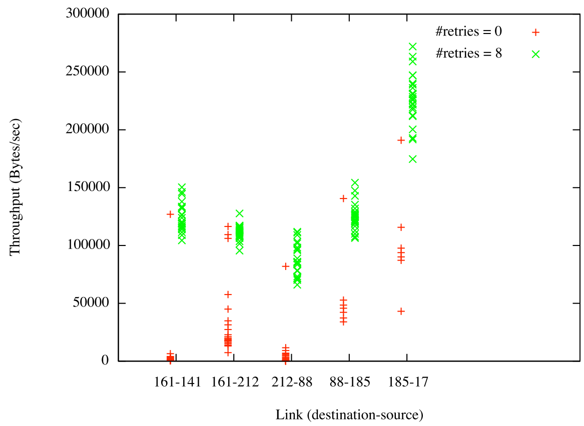

Figure 16: Effect of number of retries on throughput. Without retries the throughput decreases.

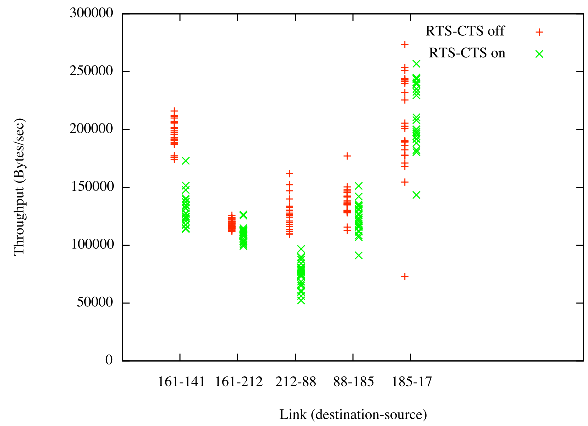

Figure 17: Effect of RTS-CTS handshake on throughput. RTS-CTS does not significantly affect the throughput.

the different values on the network performance.

5.1 Number of Retries

We measure the performance of the links as a function of number of retries at the MAC layer. We configured the number of retries to be either 0 or 8. For each link, an average throughput is computed for every hour and the experiment is conducted over a period of a day, generating 24 throughput data points. The plot in Figure 16 shows effect of the number of retries on the throughput of the links in the UNAM section. The difference difference in throughput for retries for many of the links is very large.

The throughput is consistently higher when number of retries is 8. When number of retries is 0, every lost frame results in a retransmission at the TCP layer, which increases the delay. However, when number of retries is 8, not all retransmissions request are propagated to the TCP layer. The MAC layer recovers some of the lost packets. Therefore, setting retries to 8 consistently results in higher throughput as compared to when retries is 0. We observe that the improvement in performance is higher when the distance between the nodes is higher since transit time increases with the distance. The improvement is 10 fold for the 19.6km link.

5.2 RTS-CTS Handshake

We measured the performance of the links as a function of RTS-CTS at the MAC layer. We used two configurations: enabled and disabled. For each link, an average throughput was computed for every hour and the experiment was conducted over a period of day, generating 24 throughput data points. The plot in Figure 17 shows the effect of RTS-CTS on throughput of the links in the UNAM section. As can be seen in the figure, enabling RTS-CTS decreases throughput. We believe enabling the RTS-CTS handshake does not result in any improvement because the density of the nodes is not high enough.

5.3 Transmission Bit-rate

We measured the performance of the links as a function of the transmission bit-rate at the MAC layer. We configured the bit-rate to be either one of 1/2/5.5/11Mbps or auto. The number of retries is set to 8 and RTS-CTS is disabled. We measured the throughput for each link over a period of day, and computed hourly averages generating 24 throughput data points for each link.

As seen in Figure 18, the auto bit-rate selection is not always optimal. For example, the throughput for the link between 161 and 141 is greater for the 5.5Mbps bit-rate than is it for the auto selection. The algorithm for the bit-rate selection uses the number of lost frames to adapt bit-rate. When the number of successive lost frames crosses a threshold value, the bit-rate is reduced. Similarly, the bit-rate is increased after the consecutive number of successful transmissions crosses a threshold. The threshold for decreasing the bit-rate is lower than the threshold used to increase the bit-rate and we can see that the default values for our cards are not optimal. It has been shown [4, 14] that even though there are more

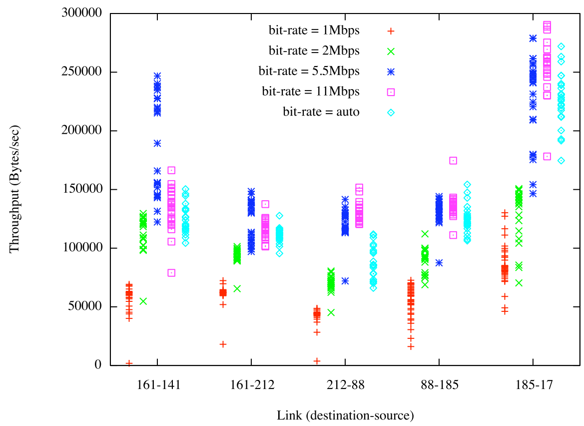

Figure 18: Effect of bit-rate selection on throughput. The auto-bit-rate selection is not optimal.

losses at a higher bit-rate, the overall throughput at the higher bit-rate can still be greater than at lower bit rates.

6 Conclusion and Future Work

Conclusion: In this paper, we present an *in vivo* study of a large scale 802.11b wireless seismic network, where the distance between nodes is on the order of 10's of kilometers. The study is carried out in diverse environments and over a long period of time. Our study is performed *in vivo*, i.e., without stopping the seismic data collection. The use of throughput instead of PDR, as a metric to gauge the performance of the network, is how we are able to be non-intrusive. Our methodology is applicable to studing the performance of an operational network without interfering with the day-to-day operation of the network. Unlike PDR which needs to be translated to understand the network capacity, our throughput-based can directly be applied in designing networks.

We studied whether there is a correlation between the throughput and (a) the distance between the nodes and (b) the SNR. We measured the variation in throughput over small and large time-scales. We analyzed the impact of MAC-layer parameters on the network throughput. We also profiled our seismic network in terms of the number of up nodes, data yield of the networks, and observed faults over a period of 8 months. This type of wireless network characterization and long-term profile is helpful in designing future wide-area wireless mesh networks.

Unlike the Roofnet study, which does not see a correlation between PDR and the distance, we see a fair correlation between throughput and distance. One of the reasons behind this difference is that while the Roofnet network was deployed using omnidirectional antennas with little line-of-sight between them, the MASE network uses directional antennas with a fair amount of line-of-sight. Similar to the observations in this paper, both Roofnet and the Houston urban network [5], observe that link quality is almost independent of the distance between sender and receiver. As seen in this paper, the degradation in the performance as the distance between the nodes increases hints to the same conclusion. However, more experiments need to be conducted to verify the conjecture. One of the possible reason could be due to the large delay spread, which occurs when the distance between the nodes is long and the signal experiences multi-path effects.

We see a high variation in throughput over small and large time scales, which is different from the DGP result that sees low variation in the PDR over small and large time scales. Although both the MASE network and the DGP testbed use directional antennas, the nodes in the MASE network are deployed at locations guided by the seismology requirements. Hence, the antennas of the nodes may be hindered occasionally by trees, structures, etc. We attribute the high variation in throughput in the MASE network to this occasional hindrance.

Future Work: Using the MASE network, we have collected data to study the Earth's upper mantle and the subduction process, coast to coast across Mexico. We have begun the work to move the MASE system to southern Peru to study a similar subduction process. The data yield of the MASE network was satisfactory for the first instance of this style of network, however it can be improved. The move to Peru will include various enhancements to the software to increase the data yield beyond the 78% yield. These enhancements include creating a set of tools for the initial site surveys to help with site selection and better mechanisms and interfaces to help understand the state of the network. We will also study whether delay spread is degrading the throughput at large distances.

In addition to deploying a network in Peru, we have also planned a seismic network system to study existing models of earthquakes using very fine sensor granularity. In particular, we are working on a rapidly deployable, heavily duty-cycled seismic network for collecting data from aftershocks. New node platforms [10] are enabling us to create lower power networks using high power radios, so that the nodes can last longer than a few days without solar power. The scale of this type of network is on the order of 100 nodes at .5 to 1km spacing, which less than that of the MASE network. The challenges would be higher contention among nodes and higher variations in the surrounding environments. The newly developed system will also be used for the structural health monitoring of tall buildings and structures.

References

[1] Middle america subduction experiment (mase). http://www.gps.caltech.edu/~clay/MASE.html.

[2] The national ecological observatory network. http://www.neoninc.org.

[3] Water and environmental research systems (waters) network. http://cleaner.ncsa.uiuc.edu/.

[4] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris. Link-level measurements from an 802.11b mesh network. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 121–132, New York, NY, USA, 2004. ACM Press.

[5] J. Camp, J. Robinson, C. Steger, and E. Knightly. Measurement driven deployment of a two-tier urban mesh access network. In *MobiSys 2006: Proceedings of the 4th international conference on Mobile systems, applications and services*, pages 96–109, New York, NY, USA, 2006. ACM Press.

[6] K. Chebrolu, B. Raman, and S. Sen. Long-distance 802.11b links: performance measurements and experience. In *MobiCom '06: Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 74–85, New York, NY, USA, 2006. ACM Press.

[7] R. Clayton. Mase: Shallow subduction in central mexico. http://www.gps.caltech.edu/~clay/MASEdir/MASEprogress_report.html, 2006.

[8] K. Fall. A delay tolerant network architecture for challenged internets. In *Proceedings of SIGCOMM*, 2003.

[9] M. Lukac, L. Girod, and D. Estrin. Disruption tolerant shell. In *Proceedings of the 2006 ACM SIGCOMM Workshop on Challenged Networks*, Pisa, IT, 2006.

[10] D. McIntire, K. Ho, B. Yip, A. Singh, W. Wu, and W. J. Kaiser. The low power energy aware processing (LEAP) embedded networked sensor system. In *IPSN '06: Proceedings of the fifth international conference on Information processing in sensor networks*, pages 449–457, New York, NY, USA, 2006. ACM Press.

[11] F. Naeim, S. Hagie, A. Alimoradi, and E. Miranda. Automated post-earthquake damage assessment and safety evaluation of instrumented buildings. Technical Report 2005-10639, John A. Martin & Associates, Inc. Research and Development Department, 2005.

[12] N. R. C. (U.S.). *Neon: Addressing the Nation's Environmental Challenges*. National Academies Press, 2004.

[13] G. Werner-Allen, K. Lorincz, J. Johnson, J. Lees, and M. Welh. Fidelity and yield in a volcano monitoring sensor network. In *7th USENIX Symposium on Operating Systems Design and Implementation 2006*, Seattle, November 2006.

[14] S. H. Y. Wong, S. Lu, H. Yang, and V. Bharghavan. Robust rate adaptation for 802.11 wireless networks. In *MobiCom '06: Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 146–157, New York, NY, USA, 2006. ACM Press.