ESD342 Roofnet Report

Derek Rayside, Jennifer Underwood, Yingxia Yang

May 2006

Annotated Table of Contents

RoofNet

1. General Overview

This is an exploratory project that reflects the diversity of its objectives:

- To study routing protocols per Jennifer's original proposal for the class project (greatly limited by the data available).
- To apply the tools/methods of the class to the project and to analyze the system from an architectural point of view
- To solve problems identified as concerns by the RoofNet team.
- To find out what's going on with the continued deployment of RoofNet in Cambridge.

These objectives converge and diverge in different ways.

2. System Description

We are studying "RoofNet". RoofNet is a somewhat ambiguous term, with several meanings.

The Many Meanings of RoofNet:

- A research group at MIT http://pdos.csail.mit.edu/roofnet/doku.php.
- The research produced by that group (eg, routing protocols, analysis of link quality, etc).
- A test deployment of that research by the research group in the area arouud MIT.
- A test deployment of that research by the research group in Tent City in Boston.
- A test deployment of that research by the city of Cambridge.

Wireless Mesh Networking:

The RoofNet deployment is a wireless mesh network. A conventional wireless network is bipartite: there are access points (gateways) and clients, and clients only connect to access points. A mesh network has two classes of nodes, but is not bipartite: the clients can all connect to each other, and route each other's packets towards their ultimate destination (which is probably a gateway). This has a number of practical advantages:

1. fewer gateways are required
2. the clients can be further from the gateways (as long as there are other clients in between to relay their packets towards the gateway)
3. the network is more robust, because each client probably has multiple viable paths to route its data on (in a conventional network, each clieut has only one path: a single liuk to its nearest gateway)

Brief Contrast with other Mesh Networks

There are a number of other wireless mesh networks. Tropos is the dominant company in the area. RoofNet is distinguished from the most other wirless mesh technologies in a number of ways:

1. Commercial mesh technologies use more conventional routing algorithms. The purpose of the RoofNet project is to explore novel routing algorithms.
2. RoofNet client nodes are entirely self-configuring. Many commercial mesh technologies require a technician to configure nodes when they are deployed.
3. Getting signal inside buildings is a major challenge for wireless mesh networks. Most commercial implementations suggest hanging more nodes on more telephone poles. The first incarnation of RoofNet technology solved this problem by running an ethernet cable from the roof into one's apartment. Landlords do not like this. The next incarnation will have two kinds of client nodes: solar powered rooftop repeaters, and small inside window-ledge repeaters. The RoofNet team is going on sabbatical to develop these window-ledge devices, and the MuniMesh team (see below) is developing the roof-top repeaters.

2.1. Stimulus, Main Actors, Stakeholders

2.1.1. RoofNet Research Team

The RoofNet research team is led by Professor Robert Morris of MIT's Computer Science and Artificial Intelligence Laboratory. Students who have worked on RoofNet include: Dan Aguayo, John Bicket, Sanjit Biswas, Ben Chambers, and Douglas De Couto.

The primary goal of the RoofNet research project is to explore novel routing algorithms for wireless mesh networks.

Most members of the RoofNet research team are going on sabbatical to a startup company named Meraki Networks, where they are working on new hardware for RoofNet nodes.

Image removed for copyright reasons.
Photo of someone holding a meraki device.

Image removed for copyright reasons.
Photo of the inside of an electrical device with the size of a playing card.

2.1.2. MuniMesh

"MuniMesh" is Kurt Keville and Bob Keyes: two "volunteer reseearchers" who are working on the technology transfer aspect of things. They work on assembling and deploying hardware for the city of Cambridge, and also on practical software engineering aspects of deploying RoofNet technology, but not on the core routing protocols. For example, they are working on making economical, physically robust, and solar-powered rooftop repeaters. The original roof-top equipment used by the RoofNet research team were regular PCs (hence relatively expensive) and required running both power and ethernet wires to the roof, which many landlords do not appreciate.

They are also writing a book on municipal wifi.

2.1.3. City of Cambridge

The main stakeholders and actors within the City of Cambridge:

- Councillor Henrietta Davis, current chair of the Cable TV, Telecommunications, and Public Utilities Committee
- Mary Hart, CIO
- Linda Turner, project manager
- Bob Coe, technical lead

Other involved parties:

- Cambridge Housing Authority
- Cambridge Health Alliance
- Museum of Science
- Harvard

2.2. Sources of Needs and Requirements

Robert Morris, the Professor leading the RoofNet project, is particularly concerned about the problem of nodes "on the periphery". Our research attempted to tackle this question in addition to analyzing the topology of the RoofNet network architecture.

2.3. System Extent (Boundary and Quantities)

The system is the RoofNet network as it existed in 2004. The boundary ends at the Gateways; there is no consideration of transition or interaction with the external WWW/internet, only the interactions internal to the wireless mesh RoofNet network. Furthermore, the aspects of the system under study are limited by the data available (please see the RoofNet Data (RawData)).

2.4. Mission Statements

Our project goals are the following:

- Analyze the effects of increasing the attempted data rate.
- Analyze and benchmark the network topological properties for the aggregate data.
- Analyze the robustness of the RoofNet architecture.
- Analyze the performance of the periphery nodes.
- Understand the current political situation in Cambridge involving RoofNet deployment.

3. System Historical Background and Evolution

RoofNet has been deployed for several years in the Central Square area of Cambridge, MA. In the course of the experiment, the RoofNet network has grown in size. For example, in 2004, the network consisted of 38 nodes and 3 Gateways. In 2005, it had grown to 50 nodes. Please see GeographicalMaps to see this evolution of the network.

The version of RoofNet deployed in the Central Square area of Cambridge, MA, consists of PCs and roof-mounted antennas. The deployment of current and future systems is moving away from this rooftop deployment and toward "small and many", similar to the concept of sensor networks. The current implementation of this morphing strategy is encompassed in the "Tent City" project described here.

The system architectural structure does not seem to be changing; deployment remains de-centralized, each node still functions as both a client and a router/repeater, and the only access to the external www/internet is through specified Gateway nodes.

The future of the RoofNet deployment in Cambridge is discussed in the CambridgePublicInternet section.

4. Assessment of System Effectiveness

Please refer to the following sections (also described in the Annotated Table of Contents section) that assess the system effectiveness:

- Effect of Increasing Attempted Data Rates
- ContrastingTopologies
- OperatorDiagnostics

The analyses performed in the above sections were conducted using the data discussed in the RoofNet Data Section (RawData).

5. Reflections and Comparisons

Please refer to the ReflectionsandComparisons page.

last edited 2006-05-16 23:28:39 by 128

RawData

1. Sources of RoofNet Data

There are two main sources of RoofNet data publicly available. The trace data for both types of data can be found here. This section dicusses the data available and motivates the use of the 2004 SIGCOMM trace data for our project. Data from 2005 was not available.

1.1. 2004 SIGCOMM Paper

The 2004 SIGCOMM paper, *Link-level Measurements from an 802.11b Mesh Network*, can be found here. The paper focuses on analyzing the patterns and causes of packet loss in the Roofnet network. This section discusses the structure and content of the trace data used in the 2004 SIGCOMM paper.

1.1.1. Nodes

The 2004 SIGCOMM data contains specific information on each of the RoofNet nodes. This information is provided in text files that identify the IP Address (or Node ID) and geographical coordinates for each node.

- **Coordinates:** The coordinates are provided in terms of latitude and longitude.
- **IP Addresses and RoofNet IDs:** The RoofNet IDs are the unique RoofNet-specific identifiers assigned to each node. The node IDs can be found by the two low-bytes of the IP Address. For example, the building NE43 Gateway IP addresses were 5.4.102.110 and 5.5.92.100. Thus, the node IDs are 26222 and 23652.

Excerpt from the Coordinates file bottom of the page here:

IP Address	Latitude	Longitude
5.3.173.178	42.363546	-071.099826
5.4.160.160	42.360150	-071.088829
5.4.160.150	42.362881	-071.110256
5.4.168.216	42.363532	-071.099663

Interestingly, there is a separate coordinates file contained within the SIGCOMM trace file. This coordinates file lists the node ID (instead of the IP Address) and geographical coordinates.

Excerpt from the Coordinates file contained within the SIGCOMM trace data:

RoofNet Node ID	Latitude	Longitude
26206	42.365494	-71.096788
23652	42.363601	-71.09108
44466	42.361125	-71.092605

1.1.2. Traffic data

The RoofNet 2004 SIGCOMM traffic data was collected in the space of a few hours over a single night. The network was separated from the Internet to ensure that no outside traffic would contaminate the data.

The data is relatively clean in the sense that it is self-contained within the RoofNet network. For this reason, it is also relatively contrived: it does not necessarily give an accurate sense of what nominal traffic levels are like. However, it does give a sense of the connectivity and topology of the network. All of our analyses in this report are based on this topological information (ie, they are not based on studies of actual traffic patterns: we are looking at a map of the territory, not video of cars on the road).

Experiments

The data is separated into 4 distinct experiments. In a given experiment, each node takes a turn sending a series of 1500-byte broadcast packets at a specified attempted data rate. All of the other nodes listen (including the Gateway nodes). Each experiment represents a different attempted data rate (1, 2, 5.5, and 11 Mbps).

Structure of the Traffic Data

The traffic data is provided in three pieces within the 2004 SIGCOMM trace data.

- **Sent Packets:** The raw data containing information on all of the packets that were sent. The data file tracks the experiment ID (corresponding to the attempted data rate), the source Node ID (who sent the packet), a unique sequence number for each packet (locally assigned), and a time stamp for when the packet was sent.
- **Received Packets:** The raw data containing information on all of the packets that were received. The data file tracks the experiment ID, the destination Node ID (who received the packet), the unique sequence number for the packet (as assigned by the sender), a time stamp for when the packet was received, and signal and noise values as measured by the 802.11 card.
- **Summaries:** The raw data containing information on each link. This data file combines information on each source and destination node pair, the experiment ID, and provides the delivery ratio as defined by the fraction of packets sent by the source that were received by the destination node. The file also notes the signal and noise average values for all received packets.

1.2. Other Data

The trace data here includes traffic data collected over the course of several months. These traces probe traffic including packets that are passing through the Gateway nodes to the rest of the internet. Although this data represents sampling of nominal traffic patterns, it is not self-contained to the RoofNet network itself, though it would provide information as to congestion patterns. Because there is much uncontrolled data, the RoofNet team suggested we not use this data.

2. Data Inconsistencies and Issues

Although the raw data provided by the RoofNet team was carefully organized and archived, we stumbled across a few challenges to implementing our desired analysis of the network architecture.

2.1. 2004 SIGCOMM Data Inconsistencies

The first challenge involved inconsistencies within the 2004 SIGCOMM data.

2.1.1. Traffic Data Inconsistent with Coordinate Data

There were eight instances in which nodes were referenced in the traffic data as source and/or destination nodes but did not appear in the Node-ID coordinate file.

2.1.2. Coordinate Data Inconsistent with Map

There were three instances in which nodes were referenced in the coordinate file but whose latitude-longitude coordinates were not consistent with the map provided in the paper.

2.1.3. Gateway Identification

The paper only makes vague references to the location of the Gateway nodes. We felt it was important to know their location in order to understand their potential impact on the architecture.

2.2. 2004 SIGCOMM Traffic Data Issues

The manner in which the traffic data was collected limited the amount and type of analysis we could meaningfully perform. We had great interest in considering the relationship between the RoofNet architecture and how it performed in terms of congestion and routing.

2.2.1. No Global Clock Synchronization

The traffic data was not synchronized. Each RoofNet node locally estimated the time a packet was sent and received. Since the clocks at each node were not synchronized, there were multiple instances of packets arriving before they were sent if a global time were assumed.

2.2.2. No Global Unique Packet Identifiers

The packet numbers were not globally assigned. Each node locally assigned unique packet identifiers. This made tracking the route packets took through the network impossible.

3. Resolution of Inconsistencies and Issues

To resolve the inconsistencies and issues discussed in Section 2, we met with members of the Roofnet team.

3.1. Resolution of Inconsistency 2.1.1.

The RoofNet team provided us with the coordinate data for 6 of the 8 inconsistent nodes.

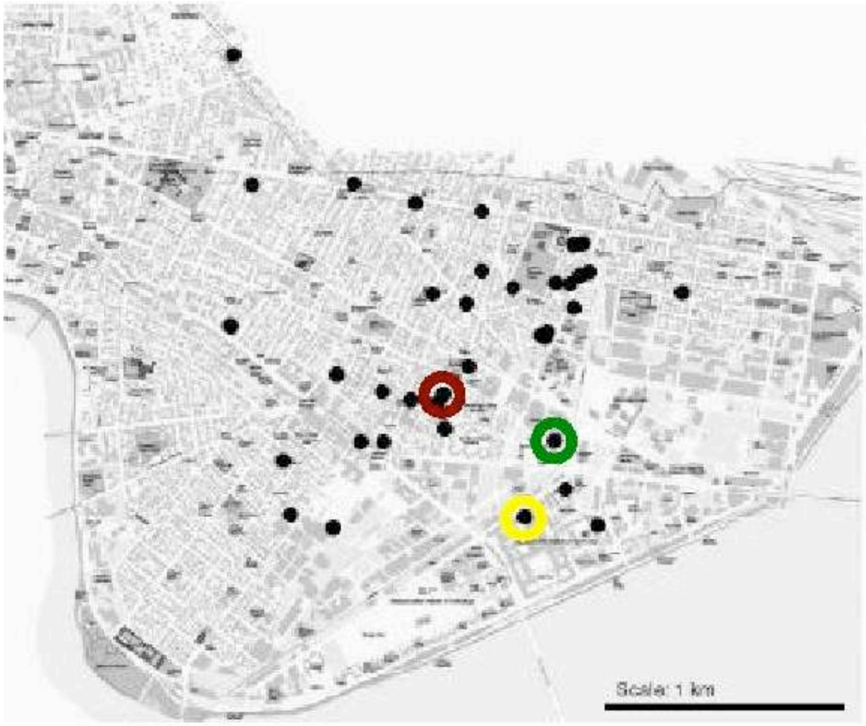
As for the other two nodes: At the time of the experiments, node 36879 did not have a separate roof-mounted antenna, but did share an apartment with 26206. They lost track of node 43220, but based on its local connections and an approximate idea of the geographical layout of the network at that time, I guessed its location.

3.2. Resolution of Inconsistency 2.1.2.

We were told that the origins of the map used in the 2004 SIGCOMM paper are lost to the mists of time. They told us to rely on the resolved data.

3.3. Resolution of Inconsistency 2.1.3.

We were given more specific Gateway information. The 2004 RoofNet map with Gateways highlighted is shown below:



LEGEND:

- **(Green) Building NE43:** Gateway nodes 26222 and 23652
- **(Yellow) Building 36:** Gateway nodes 44466/3370
- **(Red) Cherry Street:** Gateway node 26206

3.4. Resolution of Issues 2.2.1 and 2.2.2

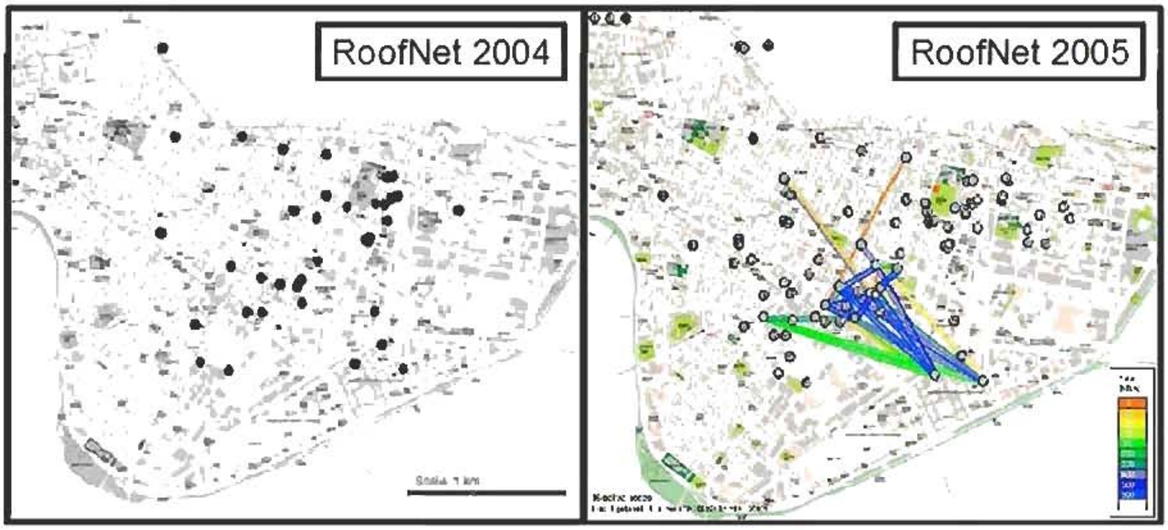
The packet/traffic data issues meant we had no real way of modeling congestion or routing performance. Any kind of traffic flow analysis would require some global knowledge of time. Thus, we could not perform congestion analysis using the 2004 SIGCOMM data. The non-unique packet identifiers was not an issue with the 2004 SIGCOMM data because of the manner in which the experiments were conducted.

last edited 2006-05-16 23:37:04 by 128

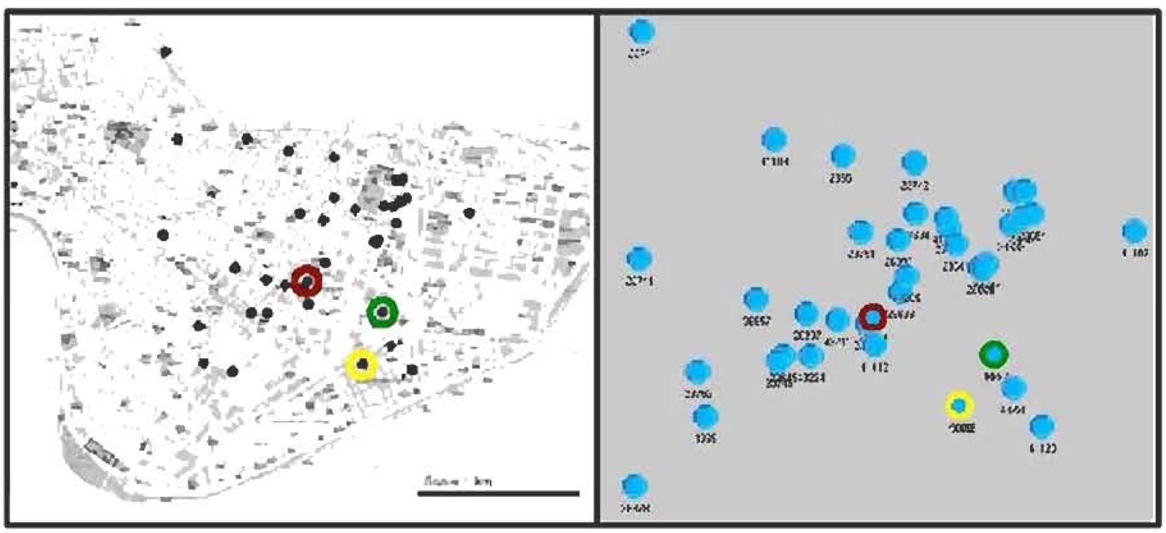
GeographicalMaps

The RoofNet network has grown since its inception. This makes it interesting as a dynamic system. The geography itself is also quite interesting since RoofNet is an unplanned network with de-centralized deployment.

Consider the evolution of the RoofNet network over the course of a year (2004 to 2005) in the maps below (courtesy of the 2004 SIGCOMM paper and the RoofNet webpage). It should be noted when viewing these maps that RoofNet was deployed to study and was not a commercial venture.



Below is a map of the 2004 Roofnet network (left, courtesy of the SIGCOMM2004 paper) as well as a map of the RoofNet coordinate data plotted in OPNET (right). The location of the Gateway nodes are highlighted in both maps. Note the differences in the map from the paper and the OPNET map using the coordinate data from the 2004 SIGCOMM raw trace data. These differences are discussed in the RoofNet Data (RawData) section of the report.



Legend:

- **(Green) Building NE43:** Gateway Nodes 26222 and 23652
- **(Yellow) Building 36:** Gateway Nodes 44466 and 3370
- **(Red) Cherry Street:** Gateway Node 26206

last edited 2006-05-16 20:10:00 by vpn-eighty-six

Effect of Increasing Attempted Data Rates

1. Effect of Increasing Attempted Data Rates

As mentioned in the RoofNet Data (RawData) section of this report, the RoofNet SIGCOMM2004 data is broken up into 4 separate experiments. In each experiment, each node attempts to send data at a specified bit rate. This protocol is in contrast to the TCP/IP protocols that adjust the bit rate real-time to compensate for congestion and poor link quality. This section discusses the analysis undertaken to understand the effect of increasing the attempted bit rates on the network topology.

1.1. Data

By Experiment: Please refer to the Roofnet Data (RawData) section of this report.

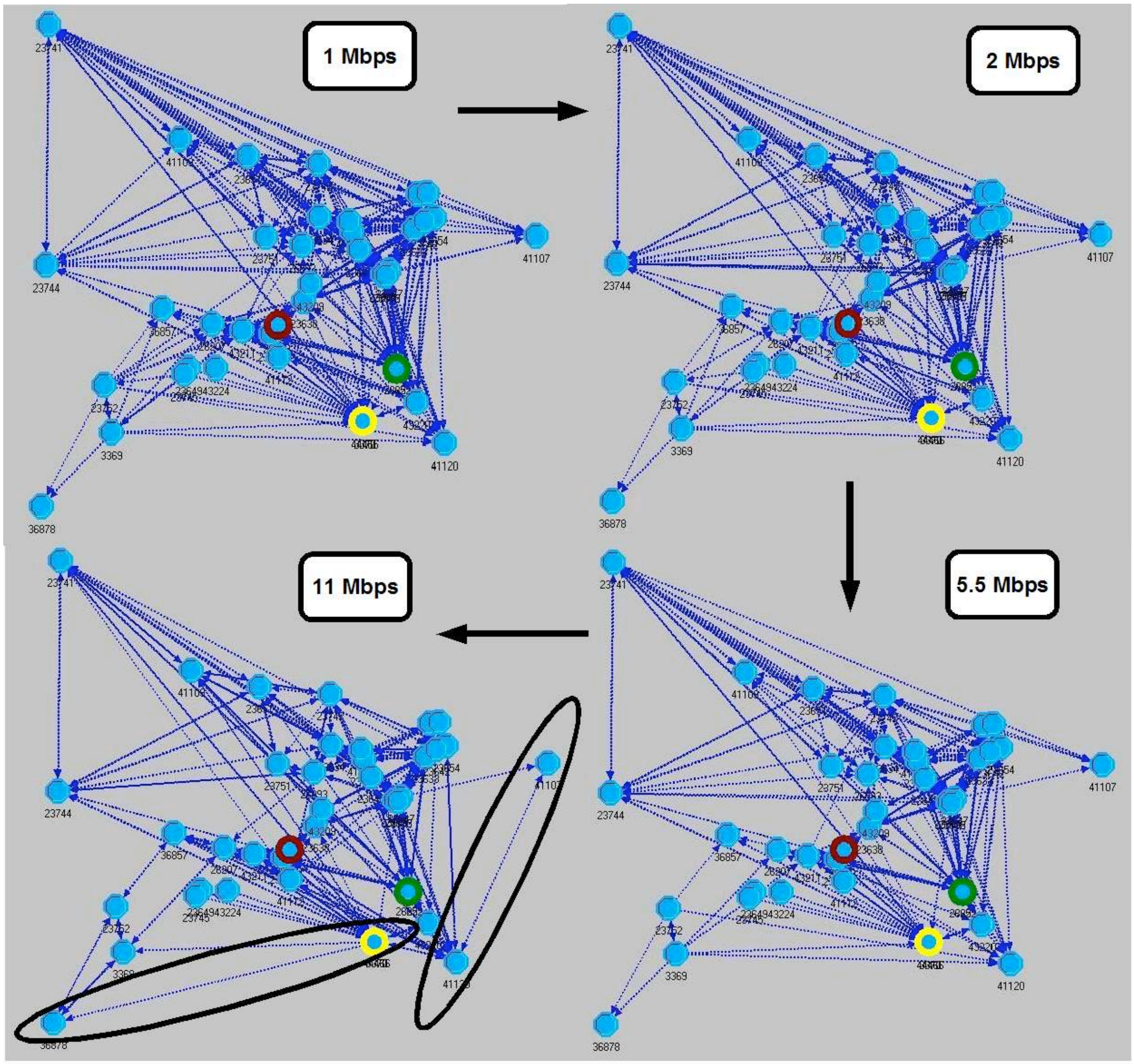
Aggregate Data: The aggregate data is a dataset constructed from the distinct experiment data for the purposes of our project. If a link between any two nodes exists at any point in time in any of the experiments, the link exists in the aggregate data. Link quality measurements are taken to be the average over all instances of the link.

1.2. Connectivity

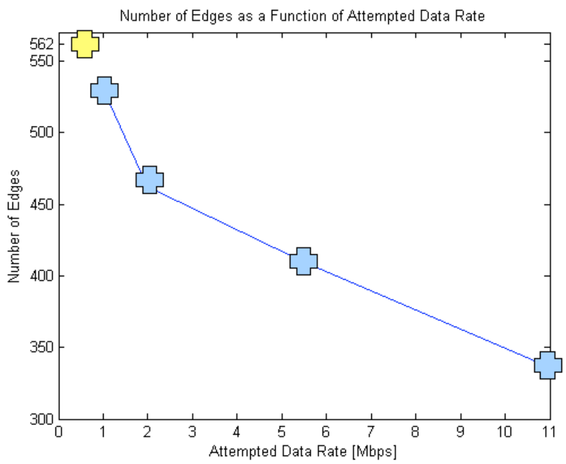
In the class, we discussed connectivity as being a metric capturing the fraction of nodes connected in a network (lecture 6). In this section, we focus our analysis of connectivity in terms of the number of edges in the network, average degree per node, and the Maximal In-degree and Out-degree. We can gain insight into the connectedness of the network topology as a whole by comparing the connectedness as a function of attempted data rates. Later sections will explore other metrics for describing connectivity.

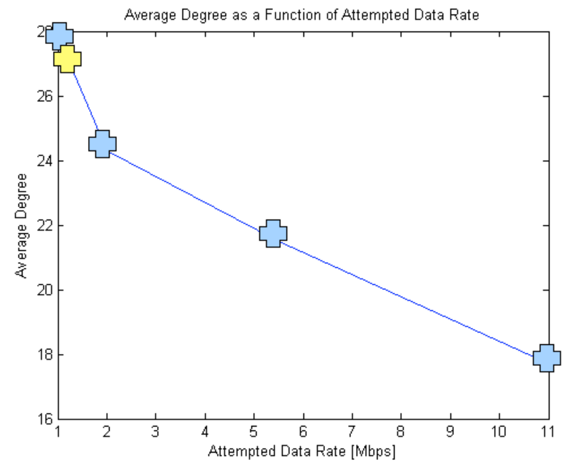
Not unexpectedly, we found that the connectivity of the RoofNet network varies as the attempted bit rates are increased. The connectivity maps for each experiment are shown below. The maps were generated by importing the 2004 SIGCOMM traffic data into OPNET. The reason for the "thinning out" of connectivity between the 1, 2, and 5.5 Mbps experiments is straightforward. Higher data rates require more energy to be successfully transmitted from one node to another. Obstacles, multi-path fade, distance, and atmostpheric phenomenon all affect the effective received energy of a signal. Thus, we expect fewer links as the data rate increases.

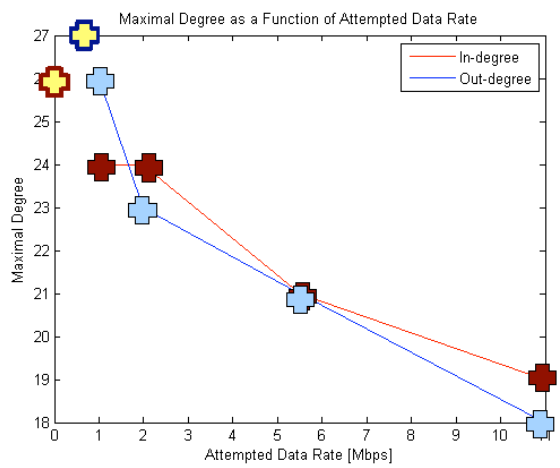
Strangely, there are two links that suddenly appear in the 11 Mbps experiment that weren't in the other experiments. These two links are circled in the 11 Mbps connectivity map below. This result is contrary to expectation given the above reasoning. However, the data was collected in a matter of a few short hours over one night. It is entirely possible that some kind of obstruction existed during the first 3 experiments that did not exist in the fourth experiment. This obstruction could be something as simple as a tree moving in the wind, a large truck temporarily parked in between the two nodes, it stopped raining, etc.



Still, the expectation that connectivity will "thin out" as data rate is increased is confirmed in the graphs below. We can see that the number of edges in the network steadily decreases as the attempted data rate is increased. Likewise, the average degree per node also steadily decreases, implying that the average number of links into and out of a given node "thins out". The differences in the Maximal In-degree and Out-degree plots imply the asymmetry of the links that is known to exist for the RoofNet network.







Attempted Data Rate	Nodes	Edges	Avg Degree	Maximal out-degree	Maximal in-degree
Aggregate	41	562	27.4	27	26
1	38	530	27.9	26	24
2	38	462	24.3	23	24
5.5	38	409	21.5	21	21
11	38	336	17.7	18	19

LEGEND:

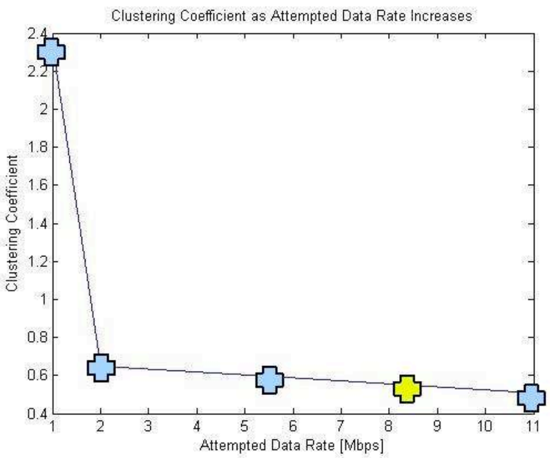
- **Blue plus sign:** symbolizes the results for each of the experiments, and Maximal Out-degree in the bottom graph.
- **Red plus sign:** symbolizes the Maximal In-degree in the bottom graph.
- **Yellow plus sign:** attempts to locate the aggregate result assuming the apparent trend continues.
- **Yellow plus sign with blue trim:** Same as Yellow plus sign but for the Maximal Out-degree.
- **Yellow plus sign with red trim:** Same as Yellow plus sign but for the Maximal In-degree.

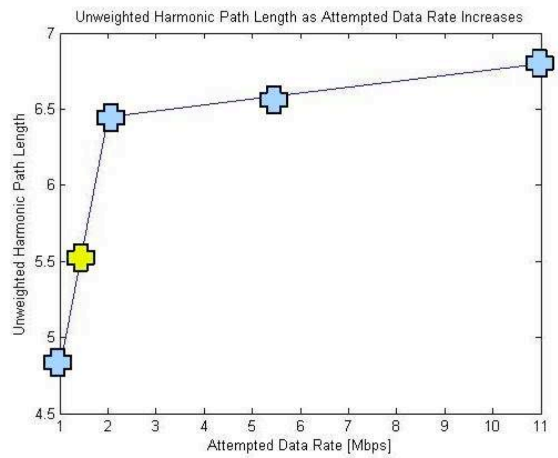
1.3. Clustering and Path Length

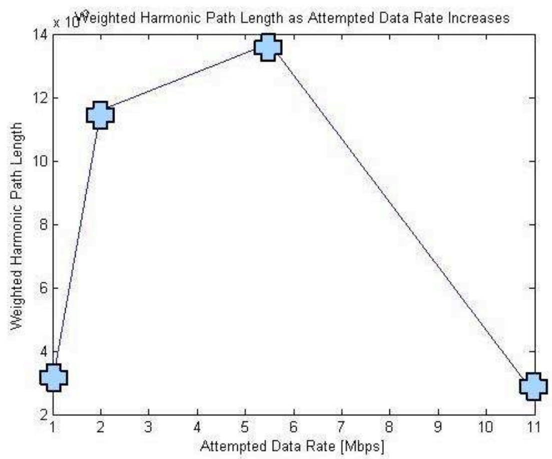
Related to connectivity are the ideas of clustering and path length. The clustering coefficient captures some knowledge about clusters of connectivity by evaluating the degree to which nodes linked to a common node are likely to have direct connectivity. Path length likewise captures some aspects of connectivity by measuring how far (in terms of number of hops, for example) a packet must travel between a source node and a destination node. The more connected the network, the shorter one would expect the path length to be.

The set of graphs below demonstrate the effect that increasing the attempted data rate seems to have on the clustering coefficient and the weighted and unweighted harmonic path lengths. The weighted path lengths account for the weight of each link on the basis of its delivery probability. Unweighted assumes that any link that exists has a weight of 1, thus making it analogous to weighting based on the number of hops to traverse the network.

The clustering coefficient drops significantly between 1 Mbps and 2 Mbps, and steadily decreases to 11 Mbps. Thus, as the attempted data rate increases, it becomes more and more unlikely that nodes linked to a common node have direct connectivity between themselves. The sudden drop between 1 Mbps and 2 Mbps could imply some kind of phase transition (the links dropped happened to be important ones, for example), though more targeted studies would have to be done to confirm this hypothesis. One would thus expect the average path length in terms of the number of hops (unweighted) to increase just as rapidly between 1 and 2 Mbps and start to level off after that (though steadily increasing). Sure enough, this is exactly what happens in the unweighted case.







Attempted Data Rate	Clustering Coefficient	Unweighted harmonic path length	Weighted harmonic path length
Aggregate	0.56250	5.59620	-
1	2.34210	4.79870	0.00300
2	0.64614	6.44820	0.01160
5.5	0.59485	6.58540	0.01370
11	0.50873	6.79820	0.00260

The weighted harmonic path length follows this trend until the transition between 5.5 Mbps and 11 Mbps when there is a sharp drop in the path length. The only effective difference between the weighted and unweighted case is that the weighted case applies more weight to links with higher delivery probabilities. Thus, the greater the path length, the greater the probability of service should be. This would imply that there is a sharp drop in the delivery probability between 5.5 Mbps and 11 Mbps. This expectation seems to be confirmed by the data in the 2004 SIGCOMM paper (see Figure 4 below).

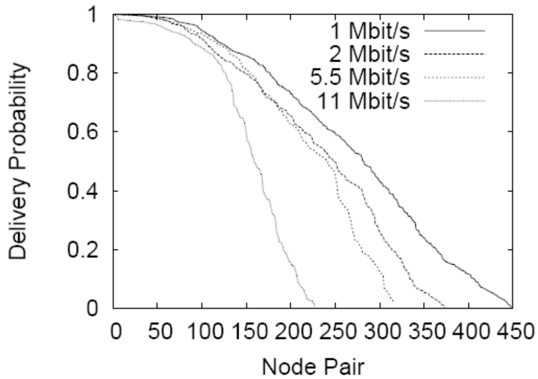
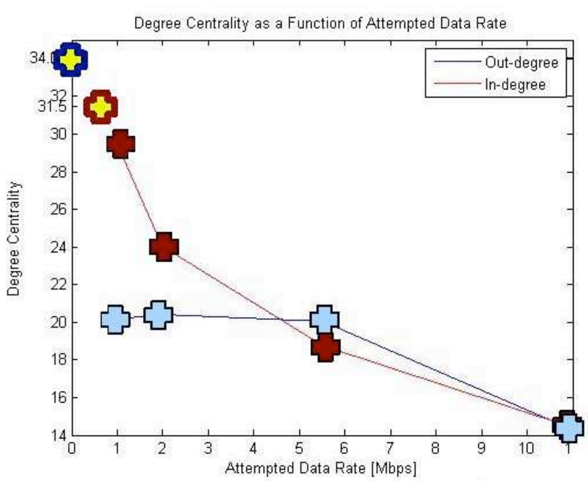


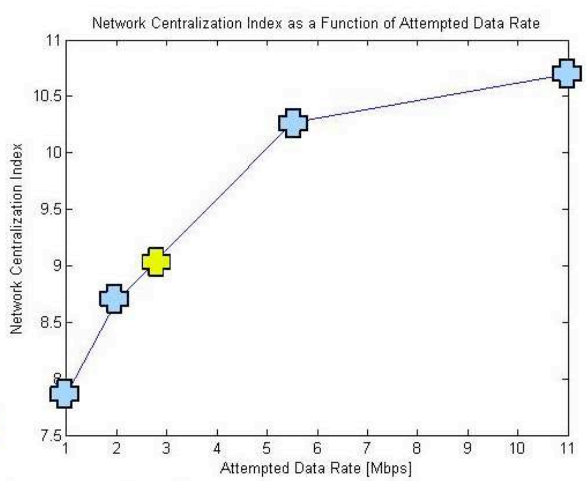
Figure 4: The distribution of link delivery probabilities for 1500-byte broadcast packets. Each point corresponds to one sender/receiver pair at a particular bit-rate. Points were restricted to pairs that managed to deliver at least one packet during the experiment. Most pairs have intermediate delivery probabilities.

1.4. Centrality

The centrality metric attempts to capture information about the amount of centralization in the network. The Degree Centrality metric defines the node that is most central as the node with the most links (Lecture 6). The Network Centralization Index (care of UCINET) measures the overall degree of centrality in the network. Ie, how much the network is controlled by nodes that are more important.

From the graphs below, it appears that the greater the attempted data rate, the more the network is controlled by more important nodes. Meanwhile, the degree centrality (both in terms of In-degree and Out-degree) decreases. This result makes sense because the more links that are dropped in the network as it "thins out" due to the increased attempted data rate, the more critical for performance certain critical paths through the network become.





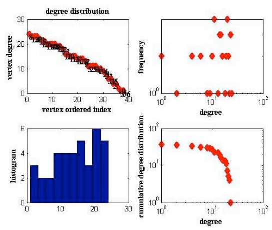
Attempted Data Rate	Degree Centrality out-degree	Degree Centrality in-degree	Network Centralization Index
Aggregate	34.06	31.5	9.19
1	20.18	29.46	7.86
2	20.4	24.05	8.69
5.5	20.04	18.76	10.27
11	14.3	14.39	10.7

1.5. Degree Distribution

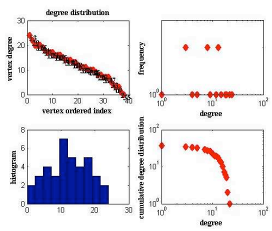
The degree distribution is a histogram of the degrees of the nodes in the network. From the graphs below, it is interesting to note that the *shape* of the cumulative degree distribution hardly changes at all as the attempted data rate increases, nor are these shapes very different from cumulative degree distribution for the aggregate data. What does happen: the graph seems to shift to the left slightly and contract ("bunch" up). Could this imply some inherent structure in the RoofNet architecture? It is difficult to say given the limited data available, but it is a curiosity since so much else seems to change significantly as the attempted data rate is increased.

There seems to be a more noticeable change in the histograms themselves. As the data rate increases, the peaks of the histogram shift left, seemingly corresponding with the shifting and contracting in the cumulative distribution.

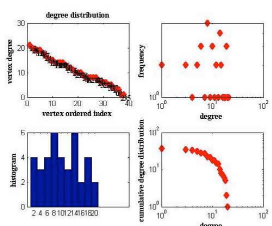
1 Mbps:



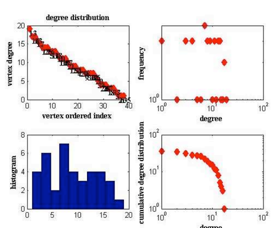
2 Mbps:



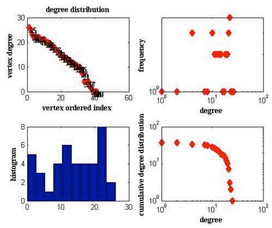
5.5 Mbps:



11 Mbps



RoofNet Asymmetrical Aggregate:



1.6. Summary

This analysis demonstrates that changing the attempted data rate in wireless mesh networks has the effect of changing the network topology. Furthermore, it seems to change the topology in largely predictable ways. Determining the extent of how this effect might be reproducible would require further analysis.

last edited 2006-05-16 20:14:45 by vpn-eighty-six