

# A Comparative Measurement Study of the Workload of Wireless Access Points in Campus Networks

Félix Hernández-Campos<sup>a</sup>    Maria Papadopouli<sup>a,b</sup>  
 Emails: {fhernand,maria}@cs.unc.edu

**Abstract**—Our goal is to perform a system-wide characterization of the workload of wireless access points (APs) in a production 802.11 infrastructure. The key issues of this study are the characterization of the traffic at each access point (AP), its modeling, and a comparison among APs of different wireless campus-wide infrastructures. Unlike most other studies, we compare two networks using similar data acquisition techniques and analysis methods. This makes the results more generally applicable. We analyzed the aggregate traffic load of APs and found that the log normality is prevalent. The distributions of the wireless received and sent traffic load for these infrastructures are similar. Furthermore, we discovered a dichotomy of APs: there are APs with the majority of clients that are uploaders and APs in which the majority of their clients are downloaders. Also, the number of non-unicast wireless packets and the percentage of roaming events is large. Finally, there is a correlation between the number of associations and traffic load in the log-log scale.

## I. INTRODUCTION

Wireless networks are increasingly being deployed and expanded in airports, universities, corporations, hospitals, residential, and other public areas to provide wireless Internet access. It is interesting to observe its evolution both in the spatial and temporal domain. While there is a rich literature characterizing traffic in wired networks (*e.g.*, [9], [8], [13], [3]), there are only a few studies available that examined and modeled wireless traffic load. Furthermore, there is no study that compares the different wireless infrastructures to generalize the models and characteristics of the traffic load. Access points (APs) are a critical element of the wireless infrastructure in campus network. The key issues of this study are the characterization of the traffic at each AP, its modeling, and a comparison among APs of two different wireless campus-wide infrastructures. Unlike most other studies, we compare two networks using similar data acquisition techniques and analysis methods. This makes the result more generally applicable.

In this paper, we study two large wireless infrastructures of the University of North Carolina at Chapel Hill (UNC) and Dartmouth College using a lightweight data acquisition methodology. The data was collected using the Simple Network Management Protocol (SNMP), the most widely available monitoring service in wireless platforms. Any AP in the market supports monitoring using SNMP, so it is important to understand how much operators and researchers can learn from SNMP data. Other types of data, such as packet or flow

level data, are generally too detailed for this purpose, and their acquisition is much more resource-intensive. This paper makes use of SNMP data for analyzing traffic characteristics, such as the total number of bytes and packets that each access points sent and received during the monitoring period. We also discuss some of the challenges of analyzing this type of data accurately.

Our study considers three dimensions of the workload an AP: number of bytes sent and received, number of packets sent and received, and number of associations and roaming operations. In addition, we have also consider how the building types (*e.g.*, academic, residential, *etc.*) affect the characteristics of the AP workloads. While previous works have partially considered some of these aspects, our focus on access points is rather unique. Furthermore, we performed system-wide characterization, rather than focusing only the most utilized areas of the studied networks. We believe this type of analysis provides a useful view of the entire utilization of a wireless network, at least from the point of the access points that form the backbone of the wireless infrastructure.

In general, we found a surprising degree of similarity in the characteristics of the UNC and Dartmouth wireless networks. Our results therefore provide strong evidence in support of the development of parsimonious workload models of campus wireless networks. This type of modeling would make it possible to develop more realistic simulations and testbed experiments.

We can summarize our contributions as follows. We analyzed the aggregate traffic load of APs across the two campus networks and found a wide range of workloads. Our analysis reveals that log normality is prevalent in both UNC and Dartmouth traces. In general, the traffic load in both wireless infrastructures is light, although there are long tails. There is no clear dependency with the type of building at which the AP is located, although some stochastic ordering is present in the tail of the distributions. An interesting dichotomy among APs is prominent in both the two infrastructures, namely, APs dominated by uploaders and APs dominated by downloaders. Specifically, we observed that as the total wireless received traffic of an AP increases, there is also an increase in its total sent traffic and at the same time a decrease in the sent to received ratio. The number of non-unicast wireless packets is substantial. Furthermore, the number of unicast received packets is strongly correlated in the log-log scale with the number of unicast sent packets. We have also studied average packet sizes. While the majority of APs send and receive packets of relative small size, a significant number of APs

<sup>a</sup>. Department of Computer Science, University of North Carolina at Chapel Hill, USA.

<sup>b</sup>. ICS-FORTH, Greece.

show rather very asymmetric packet sizes, *i.e.*, APs with an average large sent and small receive packets, and APs with small sent and large receive packets. Furthermore, we studied the distribution of the associations and roaming operations and found them quite heavy. We found a correlation of the traffic load and number of associations in the log-log scale.

Section III describes briefly the wireless infrastructure at UNC, and data acquisition process. In Section IV, we focus on analysis. Section II discusses previous related research. In Section V, we summarize our main results and discuss future work.

## II. RELATED WORK

There is only a small number of measurements studies that have examined the workload of 802.11 APs in production environments. In general, these studies have considered a wider range of issues, such as overall usage of a wireless infrastructure, and client mobility patterns, providing only a limited picture of the utilization of APs. Our work characterizes the workload of APs in a more systematic manner, and the results should have implications for the design of new wireless equipment and its evaluation.

Tang and Baker [12] used tcpdump traces and SNMP data to study a building WLAN with 12 APs and 74 users. Their only AP-specific results have to do with the variability in the maximum number users (between 3 and 12), and small number of handoffs (at most five within a five-minute period). Balazinska and Castro [2] used SNMP to characterize a much larger wireless network in three IBM buildings (177 APs). The study examined the maximum number of simultaneous users per AP (mostly between 5 and 15), total load and throughput distributions. Two interesting observation found in this paper are that offered load and number of users are weakly correlated, and that user transfer rates are dependent on the location of the AP. Balachandran *et al.* [1] performed measurements in a three-day conference setting, also focusing on the offered network load and global AP utilization. They characterized wireless users and their workload and addressed the network capacity planning problem. The overall bursty behavior and peaks and troughs are similar at all APs, though the absolute peak throughout at each AP varies. They observed that offered load is more sensitive to individual client traffic characteristics rather than just the total number of clients.

Kotz *et al.* [7], [5] studied the wireless network at Dartmouth College using syslog, SNMP, and tcpdump traces. Their first study [7] reported the distribution of average daily traffic for 451 APs, which ranged from 39 MB to more than 2 GB, and observed that maximum daily traffic was far larger than the average daily traffic. In their follow-up study [5], they reported the average number of active cards per active AP per day (2-3 in 2001, and 6-7 in 2003/2004), and average daily traffic per AP by category (2-3 times higher in 2003/2004; twice or thrice more inbound than outbound traffic). A subset of the same data (syslog messages and tcpdump traces from 31 APs in 5 buildings) was revisited by Meng *et al.* [10] for flow modeling purposes. The authors proposed a two-tier (Weibull regression) model for the arrival of flows at APs and a Weibull

model for flow residing times, and they also observed high spatial similarity within the same building. This paper makes a compelling case against Poisson modeling of wireless flows (at least for busy APs). The authors also study the modeling of flow size, and suggest that a log-normal model provides the best approximation. This is consistent with the large body of work on this topic for wired networks and file systems (*e.g.*, [11], [6], [4]).

## III. DATA ACQUISITION

The data comes from the large campus wireless networks deployed at UNC and Dartmouth. UNC's network provides coverage for 729-acre campus and a number of off-campus administrative offices. The university has 26,000 students, 3,000 faculty members, and 9,000 staff members. Undergraduate students (16,000) are required to own laptops, which are generally able to communicate using the campus wireless network. A total of 488 APs were part of the campus network at the start of our study. These APs belong to three different series of the Cisco Aironet platform: the state-of-the-art 1200 Series (269 APs), the widely deployed 350 Series (188 APs) and the older 340 Series (31 APs). The 1200s and 350s ran Cisco IOS while the 340s ran VxWorks. Dartmouth's network serves 190 buildings in a 200-acre campus. The university population includes 5,500 students and 1,200 faculty members. Laptops are also required in this institution, and almost all of them are equipped with a wireless interface. During the period of time considered in this study, the infrastructure had 557 Cisco Aironet 350 APs running VxWorks.

The data in this paper was collected using SNMP for polling every AP on campus every five minute. We collected the UNC trace using a custom data collection system, being careful to avoid the pitfalls described in [5]. The system was implemented using a non-blocking SNMP library for polling each AP precisely every five minutes in an independent manner. This eliminates any extra delays due to the slow processing of SNMP polls by some of the slower APs. The system ran in a multiprocessor system and the CPU utilization in each of the three processors we employed never exceeded 70%. The UNC trace was collected between 9:09 AM, September 29th, 2004 and 12 AM, November 25th, 2004. The monitoring system did not suffer any problems during this period.

The Dartmouth trace corresponds to the most recent dataset studied in [5]. It was collected between November 1st, 2003, and February 28, 2004, so the duration of this trace is twice the duration of the UNC one. The data was acquired using a similar approach, although the data collection system has some shortcomings that are described in section IV-A. This trace includes 6875 unique MAC addresses which were associated with one or more APs during the data collection period. This number is larger for the UNC trace, which reports on the activity of 14,712 unique MAC address. In summary, while the number of access points in both campus networks is similar, the population of wireless clients is more than twice larger for the UNC trace.

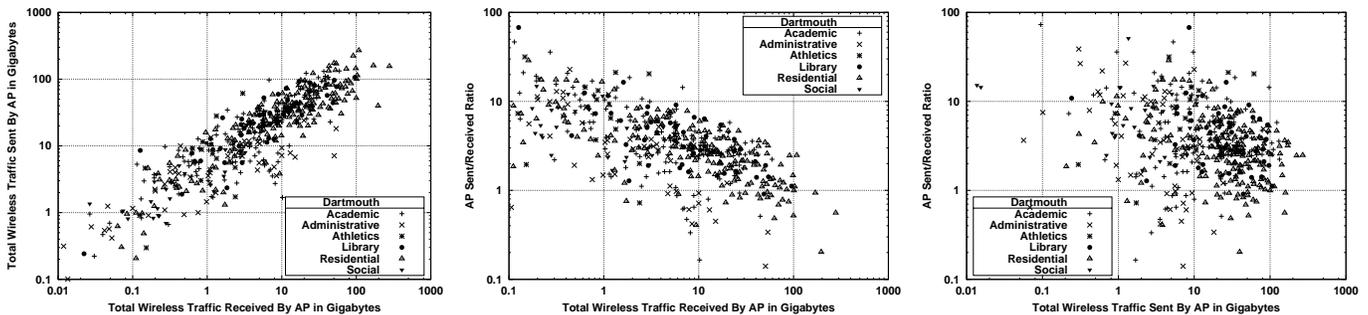


Fig. 1. Total traffic load in Dartmouth trace.

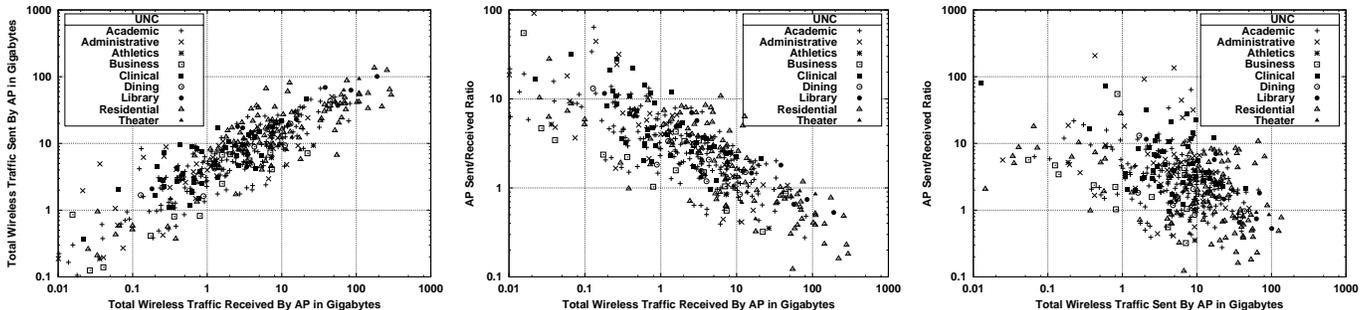


Fig. 2. Total traffic load in UNC trace.

## IV. ANALYSIS

### A. Traffic Load

The first dimension of the workload of the wireless infrastructure that we examine in this study is the total traffic load in terms of bytes. Most of the APs considered in this study made use of the 802.11b wireless LAN standard. This limited the bandwidth that could be used by an AP to communicate to its wireless client to 11 Mbps. All of the APs in the Dartmouth trace and the majority of the ones in the UNC trace used 802.11b. A higher capacity standard is 802.11a, which can reach a bandwidth of 54 Mbps. At the start of our tracing of the UNC infrastructure, 802.11a was used in 152 APs, while 802.11b was used in 335 APs. A number of the APs were upgraded during the tracing period, so at end of the trace were observed 169 using 802.11a and 311 using 802.11b.

The SNMP traces for UNC and Dartmouth consists of samples of cumulative counters. For example, one of the counters reports on the total number of bytes sent by an access point using its wireless interface since its last reboot. These cumulative counters are encoded using only 32 bits, so wrap-arounds are frequent and they must be properly handled to reconstruct cumulative values above  $2^{32}$  bytes. In this section, we first consider the total number of bytes observed by each wireless interface. Computing this quantity is straight-forward, but it requires to walk through the entire sequence of values for each counter to detect wrap-arounds. Given our sampling rate, 5 minutes, we expected that wrap-around could be detected simply by looking for values of the cumulative counters below their previous values. In practice, the analysis is far more complicated. First, every time an AP reboots, its counters are reset. It is possible to distinguish this case by examining the *up time* SNMP counter, which keeps track of the time

since the last reboot in each AP. Each reboots makes this counter decrease below the sampling interval, and this fact can be used to detect a reboot between two consecutive SNMP samples. Second, the firmware in the APs is not bug-free, and this has to be taken into account when computing the total number of bytes. One frequent problem is an off-by-one error in which the counter is decreased by one byte between two consecutive samples<sup>1</sup>. Third, the Dartmouth trace was collected using an SNMP polling system which sometimes suffered long delays to the extent that there is a significant number of samples that are in the wrong order<sup>2</sup>. This problem can be detected by checking for decreases in the up time not due to reboots. Fourth, the byte/packet counters of the Cisco Aironet 340s sometimes get reset for no apparent reason. We apply a heuristic to detect and correct this case. In summary, obtaining precise load information from this type of data requires extreme care. A single spurious wrap-around can skew the result significantly by adding Gigabytes of in-existent load.

Figures 1 and 2 provide an overview of the total traffic loads in the Dartmouth and UNC infrastructures. Since we are interested in studying the heterogeneity in the load of different APs due to their different uses, we consider here only those APs that remained operational for the majority of the tracing period. This means that the number of APs studied was 499 for Dartmouth (out of 557 present in the trace), and 447 for UNC (out of 488). The left scatter-plots in the two figures shows one symbol for each AP, comparing the total number of bytes that each access point received from its clients (x-axis) to the total number of bytes that it sent (y-axis). The

<sup>1</sup>We know this is a bug because the corresponding SNMP counters for packets do not increase accordingly.

<sup>2</sup>The researchers that collected this trace have confirmed the presence of this problem.

plots illustrate the wide range of loads in the infrastructure. Some APs had extremely light loads (a few Megabytes during months of operations) while others were used much more heavily (hundreds of Gigabytes). This is consistent for both campus wireless networks. We can also observe in both plots a clear linear trend with a positive slope. This shows that while byte loads were generally asymmetric, no extreme cases were present.

The scatter-plots in Figures 1 and 2 use different symbols for the APs located in different types of campus buildings. The building classification is based on main purpose of the building (*e.g.*, residential buildings are inhabited year-round by students). For both datasets, the majority of APs were located in academic or residential buildings. The left scatter-plots of Figure 1 and 2 show a wide range of traffic loads for each type of building. However, in both cases, the types of the most loaded APs were residential or library. Further analysis of this data using cumulative distribution functions revealed that the tails of the distributions of loads exhibit significant stochastic ordering in their tails. The tails of the distributions of loads in residential buildings are heavier than those of the academic buildings, which are themselves heavier than those in administrative, athletic and business buildings. We must note that the building type classifications, while interesting, are not completely reliable. Buildings may have areas dedicated to a different purpose and APs may serve some rooms in nearby buildings.

The middle and right plots in Figure 1 and 2 show an interesting finding. These plots use the y-axis for the ratio of bytes sent to bytes received. This quantity characterizes the symmetry in the load of APs. The smaller the ratio, the more dominated the load of the APs was by data sent from its own clients. As the two middle plots illustrate, there is a clear downward trend when the ratio of sent to received bytes is plotted against the total number of received bytes. This implies that APs with more bytes sent from their clients tend to send less data to them. We can say that these access points are dominated by *uploaders*, *i.e.*, clients that mostly serve data rather than download it from the Internet. We conjecture that this is due to peer-to-peer applications, which are fairly popular in both campus networks. The building type breakdown reveals some structure, although there is again a wide variety among the buildings of each type. Most social and dining buildings had a ratio above one, so uploading behavior was not very significant in them. Residential buildings account for a large fraction of the buildings with a ratio below 1.

The right plots in Figures 1 and 2 do not show the prominent linear trend found in the middle plots. This is most clear for the Dartmouth data. Rather than a positive trend, which would indicate that the APs with more bytes sent were dominated by downloaders, we find no trend (Dartmouth) or a slight downward trend (UNC). Since most clients are downloaders, this indicated that the total number of bytes sent by an AP increased as the number of clients increased, rather than as the clients become heavier downloaders. This is sharp contrast to the structure found for total received bytes, which increased as the APs became more dominated by uploaders.

The left plot of Figure 3 shows the cumulative distribution

functions,  $Pr\{X \leq x\}$ , for sent and received bytes. Dartmouth is heavier in both cases, as one would expect from a longer monitoring period. We also found that in general the distribution of AP aggregate loads can be well-approximated with a log-normal model. This is illustrated in the middle and left plots of Figure 3. These *quantile-quantile plots* compare the quantiles of an empirical distribution with the theoretical quantiles of a fitted model. Quantiles are represented using dots, with their theoretical value as the x-coordinate and their empirical value as their y-coordinate. In addition, the solid 45-degree line flanked by dashed confidence interval curves represents the area where the quantiles should appear in the case of a good fit. The middle plots, which correspond to the distribution of the total number of bytes sent by UNC APs, show an excellent match with the log-normal model. Note the theoretical quantiles correspond to those of a normal distribution, since the data has been transformed by taking the logarithm. Therefore, the system-wide distribution of sent bytes can be modeled accurately with a log-normal distribution. This is also true for the Dartmouth APs. On the contrary, the fit is not as good for received bytes. The highest quantiles show a systematic deviation above the upper confidence interval curve. This suggests that the empirical distribution has a tail that is significantly heavier than that of the fitted log normal model. Given the linearity of this deviation, we believe that a bimodal fit (*e.g.*, two log-normals with different parameters) would provide a good fit<sup>3</sup>. This is also true for the Dartmouth received bytes distribution. This finding is consistent with the results in Figures 1 and 2, which show two types of APs workloads (uploader vs. downloader dominated). Uploader dominated APs made the tails of the received bytes distributions heavier than captured by log-normal models. It is important to note that this conclusion cannot be extrapolated from the log-normality of flow sizes observed in previous studies<sup>4</sup> [10] (except, of course, in extreme cases of APs with only one flow in total).

The previous plots examined the total load of the APs over the tracing period. It is also interesting to study the load during shorter intervals of time. Our SNMP data was sampled every five minutes, so it seemed natural to study AP loads using such interval of time. This analysis required to handle irregularities in the SNMP polling rate. We can distinguish two types of irregularities. The first type of irregularity is the *presence of gaps* in the polling sequence. When an APs is very busy, it may not reply to the polling request from the SNMP collector. In this case, the trace shows polling samples separated by a multiple of the sampling interval. We used linear interpolation to reconstruct the values of the missing samples (*i.e.*, obtained 5-minute samples by splitting the load of a longer sample). The second type of irregularity is due the *scheduling inaccuracies* in the polling process. In the case of the UNC trace, the

<sup>3</sup>See [6] for a successful application of this approach in a similar context.

<sup>4</sup>APs handled many thousands of flows during the tracing periods. By the central limit theorem, the sum of many thousands of samples from a log-normal distribution is normally distributed (not log-normally distributed). The phenomenon we observed is due to the spacial distributions of loads across large campus networks, and not just to the distributional properties of flow sizes.

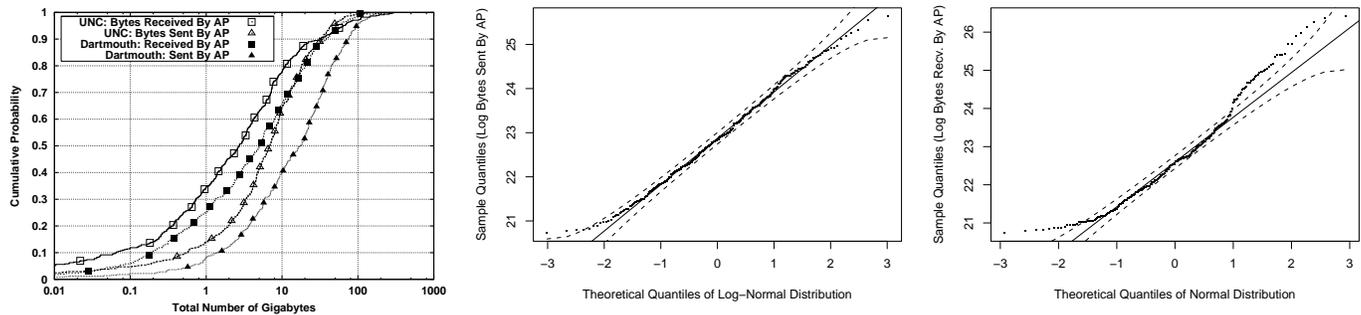


Fig. 3. Cumulative distribution functions for total traffic loads (left), and Q-Q plots for the distributions of total traffic sent (middle) and received (left) by UNC APs.

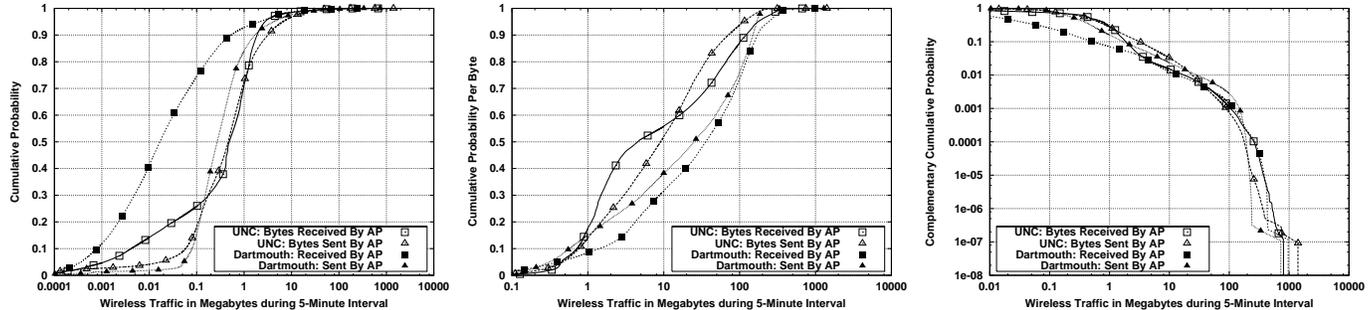


Fig. 4. Three different visualizations of the distributions 5-minute traffic loads for UNC and Dartmouth.

SNMP collector polled each access points precisely every five minutes. However, if the request was lost or the APs did not respond to it, the collector tried up to three more times, spacing each request by 5 seconds. This meant that the polling may have been late by at most 20 seconds. As in the previous case, we used linear interpolation to correct the measured values, and obtain the total number of bytes in a five minute interval. In the case of the Dartmouth trace, polling can be delayed much longer, and even be performed after the next scheduled poll. We ignored any polling not spaced at least 250 seconds from the previous poll, and made use of linear interpolation to adjust values and fill gaps.

The cumulative distribution functions shown in the left plot of Figure 4 study the load of the APs during 5-minute intervals. Note that we only report intervals with non-zero loads, and our analysis was performed for all the APs in both traces, rather than those APs that were on-line for most of the tracing period as in previous figures. This plot shows that the distributions of sent bytes are surprisingly similar for both networks. APs sent between 500 KB and 2 MB during most active intervals. In contrast, the distributions of bytes received from clients are quite different. The distribution for Dartmouth has a very light body, with less than 100 KB sent in 75% of the intervals. The UNC distribution shows two different regions: a very light one with 40% of the intervals; and a heavier one which is quite close to the distribution of sent bytes. This difference may seem at odds with the similarity between the two sites in Figures 1 and 2. However, the total bytes sent shown in those Figures are dominated by the largest values. As the middle plot in the Figure shows, the distributions of sent bytes for Dartmouth and UNC are similar when we consider the probability per byte rather than per interval. While the number

of intervals with very small loads was large, their impact was small in terms of the total bytes (*e.g.*, only 15-20% of the bytes came from the lower 95% of the intervals). Finally, the tails of these distributions are shown in the right plot of Figure 4 using a complementary cumulative distribution function,  $Pr\{X > x\}$ . The plot shows that the most utilized intervals follow similar distributions for both sent and received bytes. The sharp decrease around 1000 Megabytes is not surprising given the bandwidth limits in 802.11. The similarity, and linearity, of the curves between 1 and 100 Megabytes is more remarkable.

### B. Packet Load

A scatter-plot of the total number of unicast packets observed in UNC APs are shown in the left plot of Figure 5. The plots for Dartmouth is very similar and is not included here. The range of packets loads is very variable, with APs that handled only a few tens of thousands of packets during months of operation, while others handled tens of millions of packets. It is interesting to note that there is a far stronger correlation between the numbers of sent and received packets than between the corresponding numbers of bytes (see Figures 1 and 2). The byte plots showed a large variability around the linear trend which was due to the variable degree to which APs were dominated by more or less asymmetric communications (*e.g.*, lighter or heavier downloaders). On the contrary, when the total number of packets is examined, this heterogeneity is ameliorated by the ubiquitous use of TCP, which requires at least one acknowledgment packets for every two data packets. This severely limits the degree of asymmetry in packet load, and this observation can have implications for the design of APs and the improvement of the 802.11 protocols.

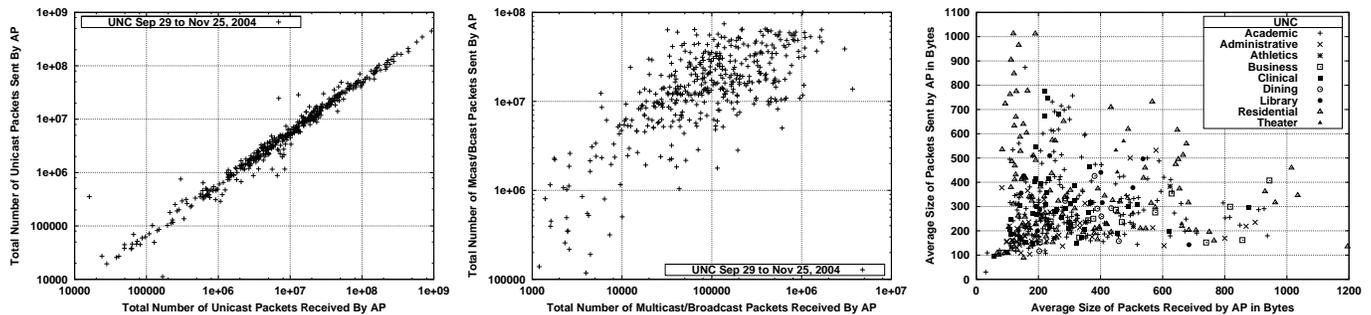


Fig. 5. Unicast (left) and non-unicast packets loads for UNC APs, and average packet sizes for UNC APs.

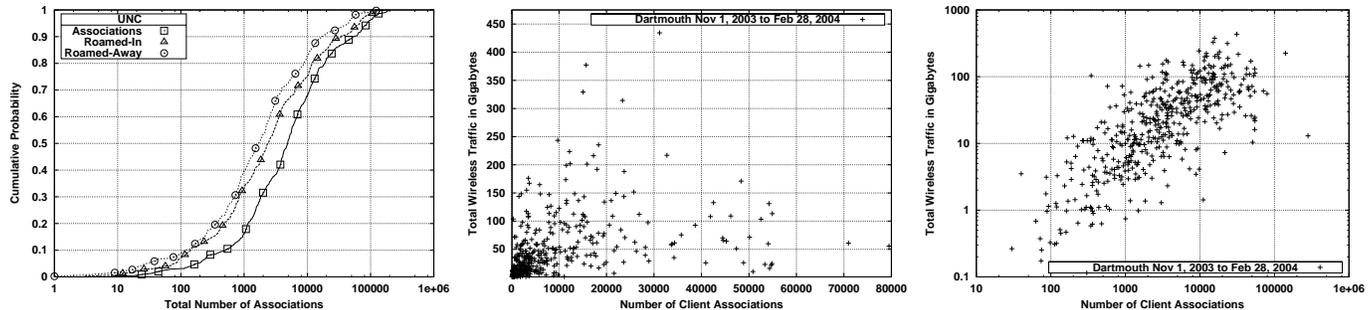


Fig. 6. Distributions of client associations for UNC APs (left) and correlation between total traffic load and number of associations for Dartmouth APs (middle and right).

The middle plot in Figure 5 shows packets load for UNC access points due to non-unicast packets. The numbers are rather large, suggesting that this type of packet are quite significant in this type of infrastructures. Unfortunately, no similar data was collected for Dartmouth. Our preliminary analysis shows that these non-unicast packets mostly came from Net-Bios and ARP.

The right plot shows the average packet sizes for the APs in the UNC campus. Most average packets sizes are below 600 bytes, although we observe some degree of *axis-hugging*, where a small average received packet size corresponds to a large average sent packet size. This confirms the previous observation that some APs are dominated by highly asymmetric client behaviors (mostly uploading or mostly downloading), which sent mostly large (data) packets in one direction and mostly small (acknowledgment) packets in the other direction. Note that this plot was created by dividing the total number of bytes by the total number of packets (both unicast and non-unicast). A similar plot for Dartmouth using total bytes divided by total unicast packets results in average packet sizes well above the Maximum Transfer Unit of 802.11b, so it is clear that the number of non-unicast packets was also rather significant for Dartmouth.

### C. Client Associations

Client association dynamics represent another important aspect of the workload of wireless APs. The SNMP counters in the Dartmouth and UNC trace include cumulative counts of the number of client that associated with each APs, number of clients that roamed into each access points, and number of clients that roamed away from each AP. The left plot in Figure 6 shows the cumulative distribution function for

these three parameters for UNC. As in previous cases, we find distributions that are quite heavy. One third of the APs received between 10,000 associations and 250,000 association during the monitored period. This represents daily averages between 170 and 4,310, which should be considered rather high. In addition, the distributions for roaming operations are also rather heavy. It is unlikely that this is purely due to client mobility, so we conjecture that associations instabilities are common, *i.e.*, clients that roamed to a nearby APs due to interferences and poor signal strength.

Previous studies [2] observed no correlation between the number of associations and the total traffic per AP. We have also studied this question and found only weak correlations for the Dartmouth and UNC traces (Pearson's correlation of 0.30 and 0.41 respectively). The middle plot of Figure 6 illustrates the absence of any clear trend in the Dartmouth data. However, when the data is examined in a log-log scale as in the right plot, it is clear that a linear upward trend is present. The correlation between the logarithm of the total bytes and the logarithm of the number of associations was 0.80 for Dartmouth and 0.74 for UNC. This means that  $\log_{10} t \approx a \times \log_{10} n + b$  where  $t$  is the total traffic and  $n$  is the number of associations. Therefore  $t \approx 10^b n^a$ , and the total amount of traffic grows very quickly with the total number of associations. In our further analysis, we have plotted (not shown here) the total traffic against the average association size, and found a upward trend. This implies that higher loads in APs with more associations are not only due a greater number of associations but also to "bigger" associations, *i.e.*, associations that transfer a greater number of bytes in average.

## V. CONCLUSIONS AND FUTURE WORK

We have studied the workload of APs in large campus networks. Our results show substantial similarities between these two environments, and open numerous avenues for further research. Our analysis of the load in bytes reveal interesting structure due to heavy uploading behavior. We intend to clarify the cause of this phenomenon by collecting and analyzing packet headers traces. This type of data would also help us to further analyze the causes of the shapes of the load distributions, and the origin of the large fraction multicast packets. Our finding of pervasive log-normality in the system-wide load of the two networks is intriguing, and we intend to study the generative process more carefully and propose more formal parametric models for the different characteristics. We also intend to study its applicability to other periods (e.g., daily loads).

We found surprisingly heavy distributions of total client associations and roaming operations. We are currently analysing an alternative source of data, syslog event messages, which should help us to estimate the fraction of these events that comes from mobility rather than infrastructural problems. Additional SNMP data that reports client information should prove useful for this analysis.

We intend to study the spatial correlations of APs and classify APs based on various parameters (e.g., traffic characteristics, number of associations, and distinct clients). Furthermore, we aim to explore the topological properties of the wireless network infrastructure.

Another ongoing effort focuses on forecasting of the traffic load at APs in various time scales. Short-term forecasting can assist in designing more energy-efficient clients. Long-term forecasting is essential for capacity planning and understanding the evolution of the wireless traffic and networks. To forecast accurately the traffic load, a very good understanding of its traffic characteristic is necessary.

This research is a part of a comparative analysis study on wireless access patterns in various environments, such as a medical center, research institute, campus, and public wireless network. We intend to analyze traces from testbeds in these environments and contrast their traffic models. We believe that understanding and forecasting the traffic of APs can have a dominant impact on the operation of wireless APs and clients and this study sets a direction for exploring further these issues.

## VI. ACKNOWLEDGMENTS

We are grateful to David Kotz and his team for sharing the Dartmouth SNMP trace with us. This work was partially supported by the IBM Corporation under an IBM Faculty Award 2003/2004 grant.

## REFERENCES

- [1] Anand Balachandran, Geoffrey Voelker, Paramvir Bahl, and Venkat Rangan. Characterizing user behavior and network performance in a public wireless lan. In *Proceedings of the ACM Sigmetrics Conference on Measurement and Modeling of Computer Systems*, 2002.
- [2] Magdalena Balazinska and Paul Castro. Characterizing mobility and network usage in a corporate wireless local-area network. In *First International Conference on Mobile Systems, Applications, and Services (iMobiSys)*, May 2003.
- [3] Mark Crovella and Azer Bestavros. Self-similarity in world wide web traffic: Evidence and possible causes. In *Proceedings of SIGMETRICS '96*, 1996.
- [4] Allen B. Downey. The structural cause of file size distributions. In *SIGMETRICS '01: Proceedings of the 2001 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 328–329. ACM Press, 2001.
- [5] Tristan Henderson, David Kotz, and Ilya Abyzov. The changing usage of a mature campuswide wireless network. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, Philadelphia, September 2004.
- [6] F'elix Hernandez-Campos, J. S. Marron, Gennady Samorodnitsky, and F. Donelson Smith. Variable heavy tails in internet traffic. *Perform. Eval.*, 58(2-3):261–284, 2004.
- [7] David Kotz and Kobby Essien. Analysis of a campus-wide wireless network. Technical Report TR2002-432, Dept. of Computer Science, Dartmouth College, September 2002.
- [8] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic. *ACM Computer Communication Review*, 25(1):202–213, 1995.
- [9] W. E. Leland, W. Willinger, M. S. Taqqu, and D. V. Wilson. Statistical analysis and stochastic modeling of self-similar datatraffic. In *Proc. 14th Int. Teletraffic Cong., 6-10*, volume 1, pages 319–328, Antibes Juan Les Pins, France, June 1994.
- [10] Xiaoqiao Meng, Starsky Wong, Yuan Yuan, and Songwu Lu. Characterizing fbws in large wireless data networks. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 174–186, Philadelphia, 2004.
- [11] Vern Paxson. Empirically-derived analytic models of wide-area TCP connections. *IEEE/ACM Transactions on Networking*, 2(4):316–336, August 1994.
- [12] Diane Tang and Mary Baker. Analysis of a local-area wireless network. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 1–10, Boston, Massachusetts, USA, August 2000.
- [13] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level. *ACM Computer Communication Review*, 25(4):100–113, October 1995.