A Framework for Wireless LAN Monitoring and Its Applications *

Jihwang Yeo, Moustafa Youssef, Ashok Agrawala

Department of Computer Science, University of Maryland
College Park, MD 20742
{jyeo, moustafa, agrawala}@cs.umd.edu

ABSTRACT

Many studies on measurement and characterization of wireless LANs (WLANs) have been performed recently. Most of these measurements have been conducted from the wired portion of the network based on wired monitoring (e.g. sniffer at some wired point) or SNMP statistics. More recently, *wireless monitoring*, the traffic measurement from a wireless vantage point, is also widely adopted in both wireless research and commercial WLAN management product development. Wireless monitoring technique can provide detailed PHY/MAC information on wireless medium. For the network diagnosis purpose (e.g. anomaly detection and security monitoring) such detailed wireless information is more useful than the information provided by SNMP or wired monitoring. In this paper we have explored various issues in implementing the wireless monitoring system for an IEEE 802.11 based wireless network. We identify the pitfalls that such system needs to be aware of, and then provide feasible solutions to avoid those pitfalls. We implement an actual wireless monitoring system and demonstrate its effectiveness by characterizing a typical computer science department WLAN traffic. Our characterization reveals rich information about the PHY/MAC layers of the IEEE 802.11 protocol such as the typical traffic mix of different frame types, their temporal characteristics and correlation with the user activities. Moreover, we identify various anomalies in protocol and security of the IEEE 802.11 MAC. Regarding the security, we identify malicious usages of WLAN, such as email worm and network scanning. Our results also show excessive retransmissions of some management frame types reducing the useful throughput of the wireless network.

Categories and Subject Descriptors: C.2.0 [Computer-Communications Networks]: Security and protection; C.2.3 [Computer-Communications Networks]: Network monitoring

General Terms: Security, Measurement, Experimentation

*This work was supported in part by the Maryland Information and Network Dynamics (MIND) Laboratory, its founding partner Fujitsu Laboratories of America, and by the Department of Defense through a University of Maryland Institute for Advanced Computer Studies (UMIACS) contract.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
WiSE'04, October 1, 2004, Philadelphia, Pennsylvania, USA.
Copyright 2004 ACM 1-58113-925-X/04/0010 ...$5.00.

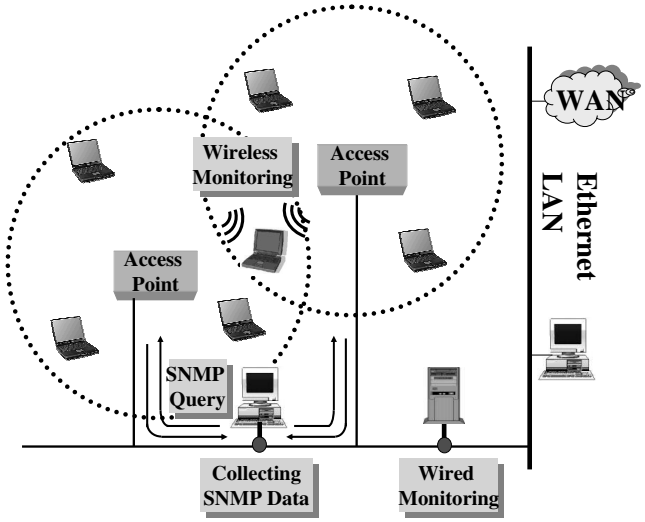

Figure 1: Monitoring Wireless Traffic: from a *wired* vantage point, a *wireless* vantage point, and SNMP statistics.

Keywords: Wireless LAN, Wireless Monitoring, Security, Anomaly, Traffic Characterization

1. INTRODUCTION

With the popularity of the IEEE 802.11 [11] based wireless networks, it has become increasingly important to understand the characteristics of the wireless traffic and the wireless medium itself. A number of measurement studies [1, 4, 7, 14, 19, 21] have examined traffic characteristics in wireless networks. In these studies, the measurements have been conducted on the wired portion of the network and/or combined with *SNMP* logs [1].

The measurements at such wired vantage points can provide accurate traffic statistics as seen in that portion of the network. However, practically they do not effectively expose the instantaneous wireless medium characteristics (PHY/MAC in the IEEE 802.11). The reasons are that SNMP exploits the summary data polled periodically (typically every 1 - 5 minutes)[1] and wired monitoring completely relies on the information observed at the wired portion.

[1]Such detailed wireless PHY/MAC information can be available with properly defined MIB (Management Information Base) and sufficiently small polling interval. However, most existing SNMP MIBs for APs [12,16,17] provide very limited visibility into MAC-level behaviors.

Instantaneous wireless PHY/MAC information is very important for security monitoring in the IEEE 802.11 wireless network. It is well known that the IEEE 802.11 WLAN has security vulnerability due to the flaws in the MAC protocol [3, 18] and basic features of wireless networks, such as open medium and mobility [26]. To correctly diagnose such security problems we need to monitor both MAC operations and mobile user activities instantaneously. Therefore, with such instantaneous wireless PHY/MAC information, security monitoring and surveillance can be effectively performed.

To capture such detailed PHY/MAC information, wireless monitoring technique can be used. Fig. 1 illustrates wired monitoring, a measurement from a wired vantage point, and wireless monitoring, a measurement from a wireless vantage point, and SNMP statistics. Recently, wireless monitoring is widely adopted in both wireless research, e.g. [20], and commercial WLAN management product development, e.g. [8, 22].

In this paper, we focus on implementing an effective wireless monitoring system and demonstrating its effectiveness in traffic characterization and network diagnosis (e.g. anomaly detection and security monitoring). Specifically, we examine the following questions: (i) In spite of the advantages for instantaneous wireless PHY/MAC information, what are the pitfalls that a wireless monitoring system needs to be aware of? (ii) How can we avoid such pitfalls to improve the capabilities of the wireless monitoring technique? (iii) How can we leverage the information that wireless monitoring provides, for WLAN traffic characterization and network diagnosis?

To answer those questions, we first conduct a number of controlled experiments to identify the pitfalls of wireless monitoring. For the identified pitfalls, we propose feasible solutions and implement them to provide a wireless monitoring system. After showing the effectiveness of the solutions, we apply our wireless monitoring system to a real WLAN traffic in computer science department in a university, over a period of two weeks. Then, we show how the monitored information can be effectively used for both WLAN traffic characterization and network diagnosis.

The contributions of this paper are as follows: (i) Our study can give the insights on how to apply the wireless monitoring technique (e.g. configuration, deployment, and data collection) for traffic characterization and network diagnosis. (ii) The characterization results can present a basis for building models and simulation tools of the 802.11 wireless networks. (iii) The anomalies identified in our study would help protocol designers to refine the protocol to remove the anomalies identified.

In the following sections, we discuss the advantages and challenges of wireless monitoring for the purpose of traffic characterization and network diagnosis.

1.1 Advantages of Wireless Monitoring

The wireless monitoring system consists of a set of devices which we call *sniffers*, to observe traffic characteristics on the wireless medium. Wireless monitoring is useful for understanding the traffic characteristics or detecting the anomaly in wireless network for the following reasons.

A wireless monitoring system can be set up and put into operation without any interference to existing infrastructure, e.g. end-hosts and network routers. In fact wireless monitoring can be performed without any interaction with the existing network, and hence is completely independent of the operational network.

More importantly, wireless monitoring exposes the characteristics on the wireless medium itself so that we can infer the PHY/MAC characteristics. Thus wireless monitoring allows us to examine physical layer header information including signal strength, noise level and data rate for individual packets. Similarly it also enables examination of the link layer headers, which include IEEE 802.11 type and control fields [11].

Physical layer information can be used to examine error rates and throughput. This is useful for developing accurate error models for the IEEE 802.11 WLANs and in site planning to determine the minimum signal strength required to achieve a certain throughput or error rate.

By analyzing the MAC layer data, we can characterize traffic according to different frame types, namely: data, control, and management frames. The collected data, combined with timestamps, can be used as accurate traces of the IEEE 802.11 link-level operations. Such traces are useful when we want to emulate the protocol or diagnose the problems of wireless networks.

1.2 Challenges of Wireless Monitoring

Despite the numerous advantages described above, wireless monitoring has the following challenges:

1. Limited capability of each sniffer: each sniffer has the limitations, e.g. on signal receiving range, disk space, processing power, etc.

2. Placement: finding the best location for each sniffer is difficult.

3. Data collection: it is difficult to collect and synchronize a large volume of data from multiple sniffers.

In this paper, we address all the above problems and propose a framework for wireless monitoring technique. However, wireless monitoring has another big challenge: *scalability*, i.e. that the cost and management overhead can be significant for the deployment and management of a large number of sniffers. In this work, we limit our work to the fixed number of sniffers for relatively small coverage area (e.g. WLAN in a single floor with less than 10 APs). Based on the promising results of this work, we are currently working towards addressing the scalability problem in more general WLAN environment.

1.3 Organization

The rest of the paper is organized as follows. In Section 2 we discuss previous works in the area of traffic characterization, network diagnosis, and security in the IEEE 802.11 WLAN. Section 3 describes the controlled experiment, the pitfalls of wireless monitoring, and the techniques to overcome them. In Section 4, we describe the results of our two-week long experiment and discuss the anomalies we discovered. Finally, we conclude the paper in Section 5 and highlight our ongoing work.

2. RELATED WORK

Several measurement and analysis studies [1, 14, 19, 21, 24] have examined traffic or error characteristics in the IEEE 802.11 WLAN. Most of the measurements have been performed on university WLAN [14, 19, 21, 24], while the work in [1] examined WLAN traffic in a conference environment.

The study of Tang and Baker [21] in the Computer Science Department building of Stanford University was one of the early studies. They examined wired monitoring traces, and SNMP logs to analyze a twelve-week trace of a local-area wireless network. In a public-area wireless network, the traces collected in well-attended ACM conference were successfully analyzed by Balachandran et al. [1]. They used SNMP logs and wired monitoring to characterize not only the patterns of WLAN usage, but also the workloads of

user arrivals and session durations with parameterized models. A significantly larger scale experiment covering a much longer duration and coverage area has been presented in the Dartmouth campus by Kotz and Essien [14]. Their analysis was based on using system logs in APs, SNMP logs and wired monitoring traces to characterize the typical usage and traffic patterns in a university WLAN. In a similar recent study, Schwab and Bunt [19] used wired monitoring and the Cisco proprietary LEAP authentication logs to characterize one-week usage and traffic patterns in a campus-wide WLAN environment.

Similar to the previous studies, our measurements are performed on typical university WLAN environment in a department network. We are interested in showing the traffic characteristics and anomalies for a typical access point in this environment. Our uniqueness comes from analyzing the wireless media using the wireless monitoring technique which gives a full view of the network spanning all the layers of the protocol stack.

In a more general wireless environment, the authors in [4,7] performed *wireless monitoring* to measure packet loss and Bit Error Rate. Their experiments were fully controlled between two wireless stations and performed on non-802.11 networks. Our work is different in being in the context of 802.11 WLANs and in performing the experiment in an actual environment with different goals.

Diagnosis for WLAN has been actively studied these days. MAC misbehaviors, such as greedy user behavior on backoff time, were examined in [10, 15]. They detected and identified the problem by either defining a new mechanism in the MAC protocol [15] or installing a detection software module in the AP [10]. Even though they did not exploit the external monitoring devices as in our work, their methodologies commonly relied on wireless monitoring technique: in [15], each wireless station (STA) performed wireless monitoring, while [10] exploited periodic wireless monitoring in the AP. Intrusion detection for mobile wireless network was addressed in [26]. They exploited an agent-based, local and cooperative intrusion detection mechanism. Wireless monitoring was also used in their work for monitoring communication activities within the radio range of each agent.

Security flaws in the IEEE 802.11 MAC have been identified and demonstrated in many literatures (e.g. [3, 5, 9, 18]). The flaws in encryption mechanisms [5, 9] and, access control and authentication [18] have been demonstrated. In [3], they identified and demonstrated two MAC vulnerabilities - namely identity vulnerability (i.e. no mechanism for verifying the correct identity) and media access vulnerability (i.e. no protection for arbitrary modification of NAV[2]). These flaws in current MAC protocol indicate that monitoring MAC operations and user activities is crucial for effectively diagnosing WLAN.

3. CONTROLLED EXPERIMENT

In this section, we present our controlled experiment. The purpose of this experiment is (i) to analyze the wireless monitoring technique in terms of its effectiveness in capturing wireless traffic and presenting precise statistics for wireless medium and (ii) to identify the pitfalls that the monitoring system needs to be aware of. Our metric for effective monitoring is the percentage of captured frames out of the frames generated by a reference application.

3.1 Methodology

3.1.1 Network Infrastructure

[2]Network Allocation Vector

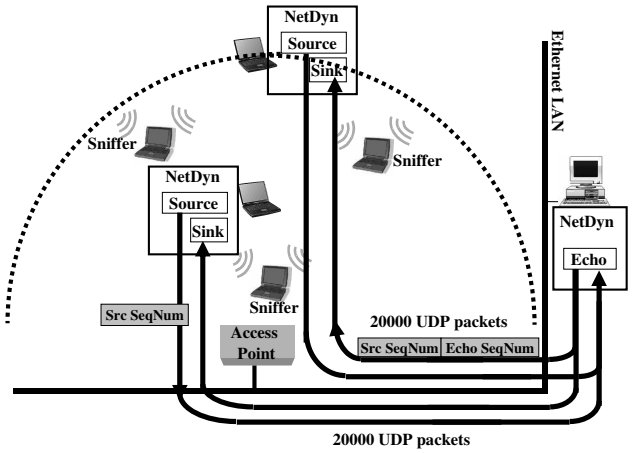

Figure 2: Controlled Experiment using NetDyn: *Source* in a wireless station sends 20000 UDP packets to wired *Echo* machine which sends them back to *Sink* in the same wireless station.

We perform our experiments in the A. V. Williams building, at University of Maryland (where the Department of Computer Science is located). The building has 58 access points installed, which belong to three different wireless networks. Each wireless network is identified with its *ESSID*. The ESSIDs of the three networks are *umd*, *cswireless* and *nist* respectively. *umd* network consists of 29 Cisco Aironet A-340 APs, and is the most widely used wireless network in the university. *cswireless* (12 Lucent APs) and *nist* (17 Prism2-based APs) networks are built by individual research groups in the department[3].

We performed our controlled experiment on a separate network that we set up specifically for this purpose with its own ESSID. Our clients were configured to associate with this AP.

3.1.2 NetDyn

To estimate the exact measurement loss, we need to use reliable application generated sequence numbers. We conducted a two-way UDP packet exchange experiments using an end-to-end traffic measurement tool, called *NetDyn* [2].

As shown in Fig. 2, NetDyn consists of three different processes, *Source*, *Echo* and *Sink*. *Source* puts a sequence number in the payload, sends the packet to *Echo*, which also adds a sequence number before forwarding it to *Sink*. In our setup, *Source* and *Sink* processes run on a wireless station, while the *Echo* process runs on a server wired to the LAN. Using the sequence numbers generated by the *Source* and *Echo* processes, we can determine which packets were lost in the path from the *Source* machine to the *Echo* machine and vice versa.

In the experiment, *Source* sends 20000 packets with the full UDP payloads (1472 bytes) to *Echo*, with 10 ms inter-packet duration (hence, at 100 packets/second). We made sure that no fragmentation occur on either side of the AP. Therefore, for each NetDyn frame on the wireless side, there is a corresponding frame on the wired side and vice versa. We use the NetDyn statistics as the baseline for comparison with the number of the frames captured by sniffers.

[3]All networks mentioned in the paper are based on the 802.11b protocol.

3.1.3 Monitoring Hardware/Software

We set up three sniffer machines to capture the wireless frames on the air. All sniffing devices use the Linux operating system with kernel version 2.4.19. We used *Ethereal* (version 0.9.6) and *libpcap* library (version 0.7) with the *orinoco_cs* driver (version 0.11b), patched to enable monitoring mode, as our sniffing software. We made use of the 'monitor mode' of the card to capture 802.11 frame information including the IEEE 802.11 header as well as physical layer header (called the *Prism2* monitor header), and higher layer protocols' information.

3.1.4 Captured Wireless Data

The wireless sniffer captures the first 256 bytes of each receiving 802.11 frame, records the complete view of the frame, i.e. PHY/MAC/LLC/IP/Above-IP information.

Prism2 monitor header is not a part of IEEE 802.11 frame header, but is generated by the firmware of the receiving card. The header includes useful PHY information, such as MAC Time, RSSI (Received Signal Strength Indication), SQ (Signal Quality), Signal strength, Noise, Signal Noise Ratio (SNR) and Data rate (in Mbps). All signal and noise information are in manufacture-specific units. However, they can be used for relative comparison.

We also capture the IEEE 802.11 MAC frame structure which incorporates the following fields: protocol version, frame type (management, data and control), Duration for Network Allocation Vector (NAV) calculation, BSS Id, Source and Destination MAC addresses, fragment, sequence number among others [11]. According to the 802.11 MAC frame type of the captured frame, we extract different information. For example, for Beacon frames, captured information include 64-bit Beacon timestamp which we use for time synchronization among multiple sniffers (Section 3.3). For Association/Disassociate and Authentication/Deauthentication frames, the information includes the reason code for such actions. We also capture higher layer protocol information, mainly for NetDyn frames.

3.1.5 Experiment Setup

We tried different scenarios for the traffic between the wireless clients and the wired server. In the rest of this section we show the results of one experiment whose configuration is shown in Fig.2. Other configurations gave comparable results. We have two wireless clients at two different locations corresponding to two different signal conditions. The "Good" client lies in an area of good AP coverage, in terms of SNR, while the "Bad" client lies in an area of bad AP coverage. We also have three wireless sniffers (T, U and V) capturing the wireless traffics between *Source*, *Sink* and the AP. Sniffer T is placed adjacent to the AP while the other two sniffers are placed as shown in Fig.3[4]. Note that the purpose of placing the sniffers in the controlled experiment was not to maximize the capture performance, but rather to study the different factors affecting the wireless monitoring performance.

3.2 Single Sniffer Statistics

We define a "From-AP" frame, as a frame transmitted by the AP to a wireless station. Similarly, we refer to a frame from the wireless station to the AP as a "To-AP" frame.

Table 1 shows the number of received packets for the NetDyn application and the percentage of MAC frames captured by the three wireless sniffers. The entries for the wireless sniffers were obtained by counting all frames with *unique sequence numbers*. In Table 1, the leftmost data column (named *NetDyn*) indicates the number of

[4]We discuss sniffers placement in Section 3.4.

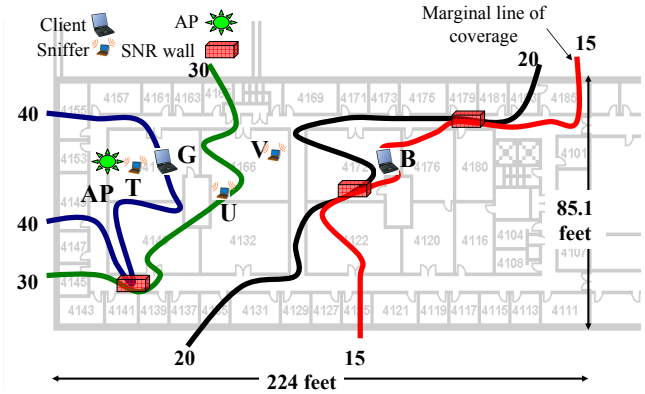

Figure 3: SNR Contour Map for controlled experiment: SNR Contour lines for 40,30,20 and 15 dB are obtained from SNR measurements. Based on the contour map, we place the wireless clients at locations G and B and place the sniffers at locations T, U, and V.

NetDyn frames correctly received by the application (out of 20000 packets). The next three columns represent the number of the frames captured by the sniffers T, U, and V respectively. Those numbers are normalized with the corresponding NetDyn number as 100%. We define the *measurement loss* to be the percentage of the frames *unobserved* by the sniffer. We can make the following observation from those four columns in the table:

- Different sniffers have different viewpoints of the wireless medium.
- The percentage of measurement loss for From-AP traffic is much less than the percentage of measurement loss for To-AP traffic. On the average, one sniffer can see 99.4% for From-AP traffic and 80.1% for To-AP traffic. The reason for that is that the AP has better hardware compared to clients, therefore the signal seen at a sniffer from an AP is stronger than the signal seen from a client. Moreover, we can always place a sniffer adjacent to an AP, whose position is fixed, while we cannot do that for wireless clients as their position is not known in advance.
- Each sniffer has a significant percentage of unobserved frames compared to NetDyn data. Even sniffer T, which was placed adjacent to the AP, encountered a severe measurement loss to observe only 73% of the total To-AP traffic.
- The *absolute* physical location of the client or the sniffer does not affect the ability of a particular sniffer to capture data from a particular wireless client. Rather, the *relative* position between the wireless client and the sniffer is the factor that affects the ability of a sniffer to capture the data from that client. For example, for the traffic originating from Bad client, sniffers U and V capture more traffic than sniffer T, because U and V are closer to Bad client than sniffer T.
- In Bad client case, the sniffers captured some frames that was not received by the NetDyn application (capture percentage $> 100\%$). This is because all sniffers are closer to the AP than Bad client which means that a frame sent by the AP will have a better SNR at the sniffer compared to Bad client. Therefore, the sniffers can capture frames that Bad client cannot capture.

Table 1: Increasing captured frames by merging multiple sniffers: Merging two or three sniffers among T, U and V significantly increases the number of observed frames.

	To-AP Wireless Traffic							
	NetDyn	*T*	*U*	*V*	*T+U*	*T+V*	*U+V*	*T+U+V*
Good	19905 (100%)	76.76%	69.00%	68.34%	76.83%	70.00%	76.84%	98.61%
Bad	18490 (100%)	69.48%	99.58%	99.73%	99.05%	100.05%	99.97%	100.13%
Total	**38395 (100%)**	**73.25%**	**83.73%**	**83.46%**	**87.54%**	**84.47%**	**87.98%**	***99.34%***
	From-AP Wireless Traffic							
Good	19247 (100%)	98.41%	97.31%	95.24%	99.37%	98.06%	99.32%	99.38%
Bad	17858 (100%)	102.04%	101.85%	102.2%	102.56%	102.43%	102.52%	102.56%
Total	**37105 (100%)**	**100.15%**	**99.5%**	**98.59%**	**100.91%**	**100.16%**	**100.86%**	***100.91%***

From these observations we can see two important pitfalls that wireless monitoring system needs to avoid for achieving a good capture percentage, i.e. a low measurement loss.

1. The capturing capability of a single sniffer is fairly limited both in terms of measurement loss and hearing (receiving) range.

2. Carefully selecting the sniffers' relative locations is important for acceptable capturing performance.

We provide the solutions for these identified pitfalls in the next sections.

3.3 Merging Multiple Sniffer Data

Our key idea for the first problem is to merge the data collected from different sniffers. Using multiple sniffers is justified not only in that it can reduce the measurement loss significantly, but also in that it can aggregate each sniffer's local view to provide a overall picture of WLAN traffic. Even if each sniffer's hardware was as good as the AP's, there is still a need for multiple sniffers in order to catch the messages that the AP missed. The main problem for merging the data from different sniffers is how to synchronize the traces when each of them is time-stamped according to the local clock of the sniffer. In this section, we describe our method for time synchronization, merging procedures and the effect of merging respectively.

3.3.1 Time Synchronization between Multiple Traces

To correctly merge multiple sniffers' data without reordering we require the time synchronization error (the difference between two timestamps of different sniffers for the same frame) to be less than *half* the minimum gap between two valid IEEE 802.11 frames. In the IEEE 802.11b protocol, the minimum gap, G_{min}, can be calculated as the 192 microsecond preamble delay plus 10 microsecond SIFS (Short Inter-Frame Space) plus 10 microsecond minimum transmission time for a MAC frame[5], to be a total of 212 microsecond.

Our approach is to use the IEEE 802.11 Beacon frames, which are generated by the AP, to be the common frames to all the sniffers. Beacon frames contain their own 64-bit absolute timestamps as measured by the AP, therefore we can uniquely identify such common beacon frames in different sniffer traces. With such n common beacon frames, we then take one of the sniffers as a reference point and use linear regression to fit the other sniffers' timestamps[6] to the reference sniffer.

Fig. 4 shows the fitting error (difference between the fitted timestamp and the reference timestamp) for the common Beacon frames over a 12.5 minutes interval. During this period, there were 5658 Beacon frames that were common to all the sniffers out of the total of the 7500 total Beacons frame that are sent at the 100 ms rate. Sniffer T was taken as the reference sniffer in this experiment. We can see that the maximum error is below 40 microseconds, well below the 106 ($= 212/2$) microseconds limit.

3.3.2 Merging Procedures

Using the obtained linear equation, we can convert the timestamp of each frame captured by each sniffer, to the reference time. To identify the duplicate frames that multiple sniffers commonly observed, we compare the header information of the frames, which are from different sniffer traces and whose converted timestamps differ by less than half the minimum gap G_{min}. After removing the duplicates, we can generate a single correctly-ordered trace from multiple sniffer traces.

3.3.3 The Effect of Merging

Table 1 shows the effect of using the merged sniffers' traces. We can see from the table that increasing the number of merged sniffers' traces from one to two to three increases the percentage of captured frames significantly from 73.25% to 84.47% to 99.34% respectively for the To-AP traffic. Notice also that the effect of merging is more significant in the case of To-AP traffic while a single sniffer near the AP (sniffer T) can almost capture all the From-AP traffic (improvement from T only to T+U+V is 0.7%).

3.4 Sniffers Placement

As noted in Section 3.2, carefully selecting the sniffers location is important to obtain an acceptable capturing performance. In this section, we describe our sniffer placement strategy in the coverage area of an AP. We make use of the observations presented in Section 3.2.

Since in the infrastructure mode of the 802.11 protocol all traffic goes through the AP, one may think that placing all sniffers near the AP should maximize the capture performance. However, our experiments showed that the capture performance of To-AP traffic is worse than that of the From-AP traffic, even for the sniffer T which was adjacent to the AP. This is due to the weak signal that reaches the sniffer from the clients compared to the strong signal

[5]For the case of Acknowledgement frame (14 bytes) transmitted at 11 Mbps.

[6]We use the MAC time of the received frame, which is available in Prism2 header in the captured frame, as the local timestamp at each machine.

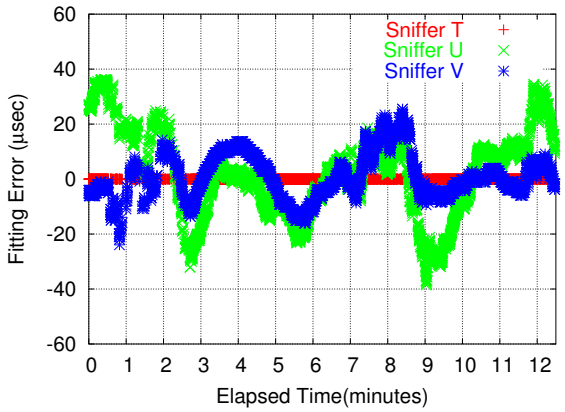

Figure 4: Fitting error with 5658 common Beacon frames (timestamp of sniffer T is the reference time).

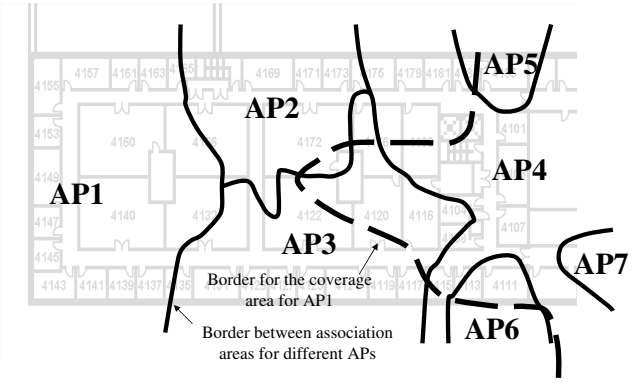

Figure 5: The Association Area for different access points. The figure also shows the coverage area for the first access point.

that reaches the same sniffer from the AP. The AP can capture the weak signal due to its better hardware and specialized processing (compared to the sniffer configuration).

Therefore, for placing the wireless sniffers, we should only place *one sniffer* adjacent to the AP to be responsible for capturing the From-AP traffic and the traffic of clients near the AP. Other sniffers should be placed as close as possible to the wireless clients.

If we assume that clients are going to be uniformly distributed over the coverage area, this translates to placing the sniffers so that they cover as much as possible from the AP coverage area. Therefore, if we have n sniffers to place, we can split the AP coverage area into n equal areas and place the sniffers in the center of mass of these areas. The key challenge here is to determine the AP coverage area. Our scheme is that at as many locations as possible we measure the SNR of Beacon frames from the target AP to draw the *contour line* as shown in Fig. 3. Therefore we can determine the AP coverage area as the 15-dB line[7] in the figure.

We can refine this strategy by noting that, in an environment where multiple APs are installed, the coverage area of an AP may be reduced to the *Association Area* of the AP. The Association Area of an AP is the area at which a client will favor this AP for association compared with other APs in the area. Note that the Association Area is a sub-area of the coverage area and that most of the traffic an AP receives comes from the associated clients (i.e. from the Association Area). Therefore, we should use the association area of an AP rather than its coverage area. Fig. 5 shows the Association Areas for different access points in the area of interest. The figure also shows the difference between the coverage area and the association area for AP_1

Another factor that needs to be taken into account is the signal condition at the sniffer location. We define an *SNR wall* as an area where the SNR contour lines are close to each other (Fig. 3). Our experiments shows that placing a sniffer near an SNR wall leads to worse capture performance compared to placing the sniffer at other places. Therefore, SNR walls should be avoided.

4. APPLICATIONS: TRAFFIC CHARACTERIZATION AND NETWORK DIAGNOSIS

We apply the wireless monitoring technique to measure and characterize actual wireless LAN traffics of a typical AP in a computer science department network. We have performed passive measurements over a period of two weeks from Monday, February 9 to Sunday, February 22 to observe the wireless PHY/MAC characteristics in the fourth floor of the A.V. Williams building of Computer Science Department on the campus of University of Maryland. In this experiment, we aim (i) to examine typical characteristics of MAC traffic in academic research environment and (ii) to diagnose various anomalies in protocol and security of the IEEE 802.11 MAC using wireless monitoring technique.

4.1 Methodology

4.1.1 Target Traffic

In the fourth floor of A.V. Williams building, we have three channel-6 APs, three channel-1 APs and one channel-1 AP. Channel 6 is the most widely used in the fourth floor, therefore we choose channel 6 as our target channel. We choose one of the channel-6 APs in the fourth floor as *our target AP*.

4.1.2 Setup and Placement

The setups for H/W and S/W in three sniffers are exactly the same as in controlled experiment in Section 3. We also followed our strategy for sniffer placement as discussed in the Section 3.4. Sniffer *T* is the sniffer placed adjacent to the target AP.

4.2 Results

We will present our results under three categories: *MAC Traffic*, *MAC Frame Types*, and *User Activity*. We also summarize the anomalies we discovered and discuss the related security issues in the last section.

4.2.1 MAC Traffic

Fig. 6 shows the daily traffic over the two weeks. We obtain MAC type and size information from each frame's MAC header in the traces. We separately present traffic for IEEE 802.11 Data frames and that for the IEEE 802.11 Management frames (e.g. Beacon frames).[8]

We notice that there was almost no user activity on the weekend of Feb. 14 and Feb. 15. This weekend represents the weekend for

[7]We can determine this threshold (in dB) by considering the reasonable percentage of captured Beacons out of normal Beacon rate, e.g. 10 per second.

[8]Since the IEEE 802.11 control frames (e.g. Acknowledgement) have no *BSSID*, i.e. MAC address of AP, we do not present the results for the control frames in this section.

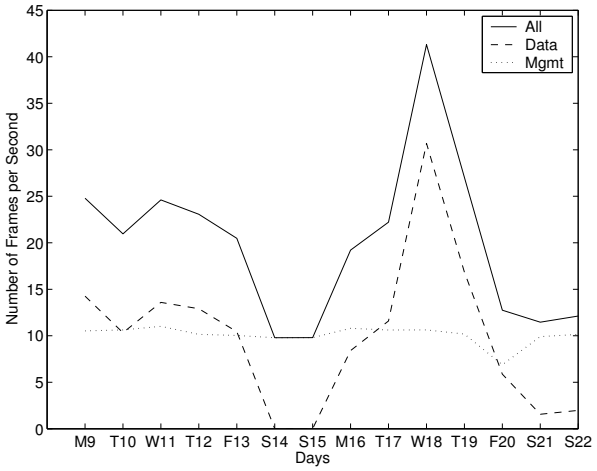

Figure 6: [MAC Traffic] Number of MAC frames per seconds, averaged daily, over two weeks: *All* traffic of the target AP is the sum of *MAC Data* traffic and *Mgmt* (Management) traffic.

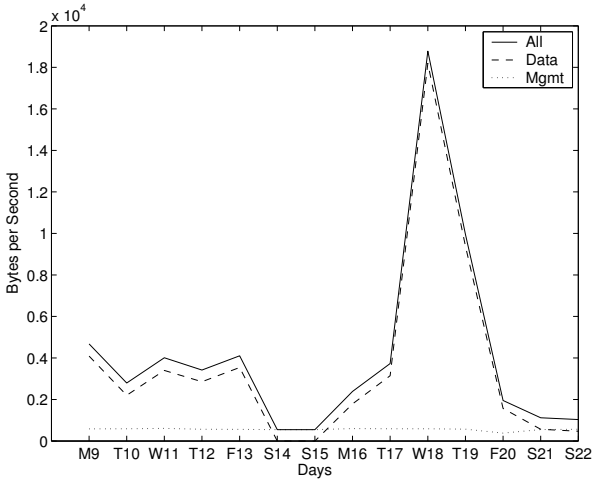

Figure 7: [MAC Traffic] Traffic volume per seconds. Daily averaged values are shown over two weeks: *All* traffic of the target AP is the sum of *MAC Data* traffic and *Mgmt* (Management) traffic.

Table 2: Abbreviation for the IEEE 802.11 Types

Abb.	*802.11 Types*
ProbeReq	Probe Request
ProbeRes	Probe Response
PowerSave	Power Save Poll
AsscReq	Association Request
AsscRes	Association Response
ReAsscReq	Reassociation Request
ReAsscRes	Reassociation Response

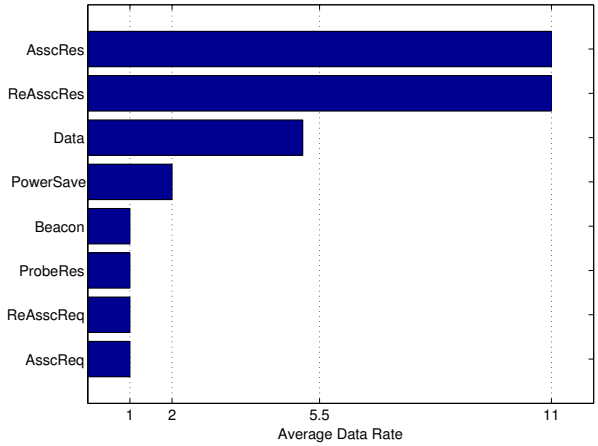

Figure 8: [MAC Frame Types] Average data rate per MAC Type.

Valentine's Day. February 16 was the holiday for President's Day. However, the university was open on that day.

Typically traffic for the IEEE 802.11 Management frames is constant over the two weeks period. On Friday, February 20, disk space had been full for 8 hours therefore we have smaller number of frames than normal days (losing about one third of the normal management traffic volume).

We observe a sudden spike of traffic on Wednesday, February 18, which is three times larger than normal days. Carefully examined, we found that 40% of MAC Data traffic consists of *IMAP* (Internet Message Access Protocol) frames. IMAP protocol is used when client STA accesses electronic mail or bulletin board messages that are kept on a (possibly shared) mail server [13]. This abnormal spike of email traffic is due to email worm *W32.Netsky.B@mm* that was spreading on the web on Feb. 18 [6].

4.2.2 MAC Frame Types

In this section we show the results of per frame-type statistics over two weeks. For each type of frames we observed, we show the average data rates per frame and average retransmissions per frame, respectively. We obtain this information from the 256 byte MAC header of the IEEE 802.11 frames and the *Prism2* header which is generated per frame by the sniffer device driver (PHY information). We use the abbreviation in Table 2 to denote the long type names.

In Fig. 8, we have two observations:

1. AsscRes and ReAsscRes are usually transmitted using the highest data rate, i.e. 11 Mbps, while the corresponding Request frames use the lowest data rate, i.e. 1 Mbps. This is not expected as the AP should respond with a data rate close

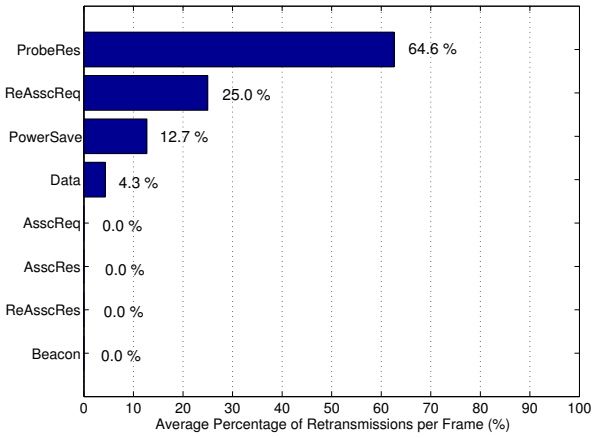

Figure 9: [MAC Frame Types] Average number of retransmissions per MAC Type.

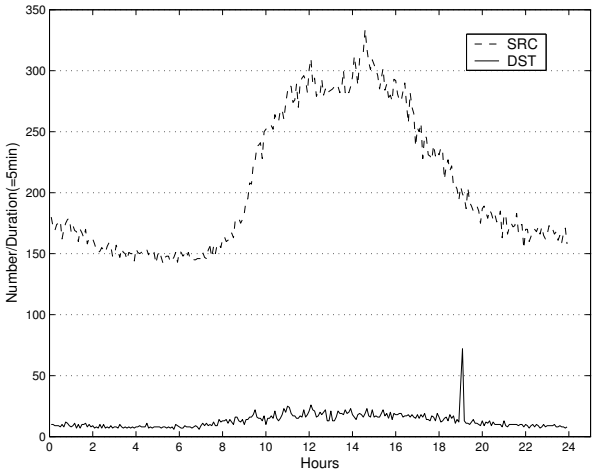

Figure 10: [User Activity] Number of Unique MAC Addresses per 5 minute interval in From-AP Traffic on Feb. 10.

to the data rate of the request to enhance the SNR at the requesting client.

2. The average data rate for Data frames is 5.1 Mbps. This strongly indicates that multiple data rates are used.

Fig. 9 shows the average number of retransmissions per frame. We calculate the numbers by dividing the number of retransmitted frames (whose MAC Retry bit is set to 1) by the number of non-retransmitted frames (whose MAC Retry bit is set to 0). Therefore, an average number of retransmissions of one indicates that each frame is retransmitted one time on the average.

We find that unexpectedly, ProbeRes, ReAsscReq and Power-Save frames have a very high number of retransmissions on the average.

We give the following possible explanations for each case:

- *ProbeRes*: Due to the overlap between different channels in An 802.11 WLAN, an AP can hear on up to 8 other channels than its assigned channel [25]. For example, suppose that a client sends a ProbeReq frame on channel 2. An AP on channel 6 can hear this ProbeReq (due to the channel overlap) and responds to the sender. When the client receives the ProbeRes frame, it knows that this is an incorrect response (since the ProbeRes frame contains the channel number it was sent on) and drops the frame. As a result, the AP on channel 6 will not receive an Acknowledgement frame and will retransmit the ProbeRes frame.

- *ReAsscReq*: Although the client sends a request with a low data rate (indicating a poor signal condition), the standard does not force the implementation to respond with a specific data rate. The AP sees from the ReAsscReq that the client can support up to 11 Mbps and sends the ReAsscRes with that rate[9]. Unfortunately, this message does not reach the client due to the poor signal conditions at the client side. This can be confirmed in the average data rate per frame types in Fig. 8.

- *Power-Save Poll*: When a STA wakes up, it sends a Power-Save Poll message to the AP asking for the buffered frames. The AP may have its *NAV* set indicating that the medium is busy, so it cannot respond to the poll message. Therefore, the station does not get a reply and retransmits the Power-Save Poll.

[9]The ReAsscRes frame acts as the Acknowledgement in this case.

We believe that these high average retransmission for this frame types represents anomalies in either the protocol design or implementation.

4.2.3 User Activity

For monitoring instantaneous user activity, we compare the number of unique source MAC addresses and the number of unique destination MAC addresses during each 5 minutes, as shown in Fig. 10. We obtained the MAC addresses information from MAC header of each frame.

We can observe that around at 19:00 on February 10 the number of unique destination MAC addresses sharply increased while the number of unique source MAC addresses did not change much. Careful examination of the trace reveals that during that 5 minute period a source station on the wired network sent ICMP ping messages to 55 unique (wireless) destination MAC addresses which are sequentially generated. This network scanning technique, namely *ping sweep*, is known to be potentially malicious.

4.3 802.11 Anomaly and Security Monitoring

With the instantaneous PHY/MAC information available through wireless monitoring, we discovered several anomalies in protocol and security of the IEEE 802.11 MAC:

1. Some management frames, e.g. association response and re-association response frames, are transmitted at the highest data rate which does not correspond to the client SNR conditions (Fig. 8). This leads to excessive retransmissions of these management frames.

2. We observe significant number of retransmissions of the IEEE 802.11 Management frames. Those frames include Probe Response (64%), Reassociation Request (25%) and Power-Save Poll (13%). These retransmissions lead to the unnecessary waste of the scarce wireless capacity. We believe the reason for such retransmissions to be the incomplete specification of current MAC protocol. To prevent such anomalies, MAC protocol standards need to specify in more detail the

frame exchange sequences and need to consider various conditions on PHY layer, e.g. data rate, signal strength, etc.

3. By monitoring the traffic and user activity, we observe some malicious usages of WLAN, such as email worm and network scanning technique.

The anomalies in MAC protocol, e.g. severe retransmission of management frames, can be exploited by malicious users for security attack. For example, we observed up to 5000 Probe Response frames (from only *one* AP, including retransmissions) during 5 minutes, which is equivalent to about 6.6 kbps. This indicates that Probe Request flooding attack can have more significant impacts on WLAN's bandwidth than are intended, due to the retransmission of Probe Responses.

We showed monitoring user activity with MAC addresses in Section 4.2.3. Other MAC information, such as MAC sequence numbers and ESSID, can also be used for monitoring suspicious user activity. In [23], they analyzed MAC sequence numbers to detect so-called fake AP and MAC address spoofing. Using multiple sniffers, our wireless monitoring framework provides accurate MAC information instantaneously on the whole MAC traffic over one or multiple BSS. Therefore, various techniques using MAC information for security monitoring can be effectively applied in our framework.

5. CONCLUSIONS

In this paper, we address two problems: wireless monitoring technique and its applications in MAC traffic characterization and network diagnosis. We first identify the pitfalls of wireless monitoring and provide two feasible solutions, namely merging multiple sniffers and their placement. Then, we apply those techniques to academic research WLAN over two week for MAC traffic characterization and network diagnosis.

Our experimental results reveal not only typical WLAN traffic characteristics but also the anomalies in the MAC protocol (e.g. severe retransmission of some Management frames) and security (e.g. internet worm and ping sweep).

We expect that the techniques and experiences in this paper can be well applied to security monitoring in the IEEE 802.11 WLAN. Because the IEEE 802.11 MAC protocol is highly vulnerable to security threats, wireless monitoring on MAC operations and user activity and the diagnosis based on such data are crucial for securing WLAN. In Section 4.2.3, we show an example of detecting a malicious WLAN usage based on the monitored user activity.

Our on-going work is centered on automating the diagnosis processes for detecting various anomalies (e.g. in terms of performance, security, and availability). We are applying multivariate anomaly detection technique to different measurement sets (e.g. SNMP and wireless monitoring) for automatic diagnosis of various anomalies.

Acknowledgement

We are specially grateful to Bao Trinh for the helps in carrying out most of the experiments in this work. We also want to thank Nikhil Purushe for the helps in SNR measurement in Section 3.4. Finally, we thank the anonymous reviewers for their valuable comments and suggestions.

6. REFERENCES

[1] A. Balachandran, G.M. Voelker, P. Bahl and V. Rangan. Characterizing User Behavior and Network Performance in a Public Wireless LAN In *Proc. ACM SIGMETRICS 2002*, Marina Del Rey, CA, June 2002.

[2] S. Banerjee and A. Agrawala. Estimating Available Capacity of a Network Connection. In *Proceedings of IEEE International Conference on Networks*, September 2001.

[3] J. Bellardo and S. Savage. 802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions. In *Proceedings of the USENIX Security Symposium*, Washington D.C., August 2003.

[4] B.J. Bennington and C.R. Bartel. Wireless Andrew: Experience building a high speed, campus-wide wireless data network. In *Proceedings of MOBICOM*, September 1997.

[5] N. Borisov, I. Goldberg and D. Wagner. Intercepting Mobile Communications: The Insecurity of 802.11. In *Annual International Conference on Mobile Computing And Networking*, Rome, Italy, July 2001.

[6] Computer Associates. Virus Information Center (Win32.Netsky.B Virus) http://www3.ca.com/virusinfo/virus.aspx?ID=38332

[7] D. Eckardt and P. Steenkiste. Measurement and Analysis of the Error Characteristics of an In-Building Wireless Network. In *Proceedings of SIGCOMM*, August 1996.

[8] Enterprise Wireless LAN Security and WLAN Monitoring. http://www.airdefense.net/

[9] S. Fluhrer, I. Mantin and A. Shamir. Weaknesses in the Key Scheduling Algorithm of RC4. In *Lecture Notes in Computer Science*, 2259, 2001.

[10] M. Raya, J-P. Hubaux and I. Aad. DOMINO A System to Detect Greedy Behavior in IEEE 802.11 Hotspots. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, Boston, MA, June 2004.

[11] IEEE Computer Society LAN MAN Standards Committee. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. In *IEEE Std 802.11-1999*, 1999.

[12] IEEE Computer Society LAN MAN Standards Committee. IEEE 802.11 Management Information Base In *IEEE Std 802.11-1999*, 1999.

[13] THE IMAP Connection. http://www.imap.org/

[14] D. Kotz and K. Essien. Analysis of a Campus-wide Wireless Network. In *Proc. the Eighth Annual International Conference on Mobile Computing and Networking (MOBICOM 2002)*, Atlanta, GA, September 2002.

[15] P. Kyasanur and N. Vaidya. Selfish MAC Layer Misbehavior in Wireless Networks. In *IEEE Transactions on Mobile Computing*, April, 2004.

[16] K. McCloghrie and M. Rose. RFC 1066 - Management Information Base for Network Management of TCP/IP-based Internets. TWG, August 1988.

[17] K. McCloghrie and M. Rose. RFC 1213 - Management Information Base for Network Management of TCP/IP-based Internets: MIB-II. TWG, March 1991.

[18] A. Mishra and W.A. Arbaugh. An Initial Security Analysis of the IEEE 802.1X Standard. CS-TR 4328, Department of Computer Science, University of Maryland, College Park, December 2002.

[19] D. Schwab and R Bunt. Characterising the Use of a Campus Wireless Network In *Proc. INFOCOM 2004*, Hong Kong, China, March 2004.

[20] M. Shin, A. Mishra, and W. Arbaugh. Improving the Latency of 802.11 Hand-offs using Neighbor Graphs. In *Proc. INFOCOM 2004*, Hong Kong, China, March 2004.

[21] D. Tang and M. Baker. Analysis of a Local-Area Wireless Network In *Proc. the Sixth Annual International Conference on Mobile Computing and Networking (MOBICOM 2000)*, Boston, MA, August 2000.

[22] Wireless Security Auditor (WSA). http://www.research.ibm.com/gsal/wsa/

[23] J. Wright. Detecting Wireless LAN MAC Address Spoofing. http://home.jwu.edu/jwright/papers/wlan-mac-spoof.pdf

[24] J. Yeo, S. Banerjee and A. Agrawala. Measuring traffic on the wireless medium: experience and pitfalls. CS-TR 4421, Department of Computer Science, University of Maryland, College Park, December 2002.

[25] M. Youssef, L. Shahamatdar, and A. Agrawala. The IEEE 802.11 Active Probing Mechanism: Analysis and Enhancements. CS-TR-4613, Department of Computer Science, University of Maryland, College Park, August 2004.

[26] Y. Zhang, W. Lee and Y. Huang. Intrusion Detection Techniques for Mobile Wireless Networks. In *Wireless Networks*, Vol. 9(5), pp. 545-556, 2003.