The Case for Heterogeneous WLAN Environments for Converged Networks

Markus Tauber, Saleem N. Bhatti
University of St Andrews, UK
{markus,saleem}@cs.st-andrews.ac.uk

Nikolay Melnikov, Jürgen Schönwälder
Jacobs University Bremen, Germany
{n.melnikov,j.schoenwaelder}@jacobs-university.de

***Abstract*—We demonstrate that in a future converged network scenario, it may be beneficial to allow selection of the used IEEE 802.11 variant based on application requirements. We have evaluated a *performance envelope* derived from testbed experiments for individual IEEE 802.11 variants and compare these with the traffic patterns from a large campus network. To demonstrate our approach, we analyse traces from the campus network from the University of Twente, comprising $\sim$5000 users. From our comparison we find that specific IEEE 802.11 variants (e.g. 802.11g or 802.11n) may be better suited to specific applications, such as video streaming, rather than using a single WLAN standard for all traffic.**

Wireless networks (WLANs) are increasingly used in home and office environments, especially in a *Next Generation Converged Networking* context, in which we aim to move services into IP networks – which may rely heavily on WLAN for the access network. However, applications, used in such a context, are not often designed or built for use specifically over WLAN. There is great value in determining the performance of such applications on real networks from studies of network traces [1]. In our discussion, we use the term *group of applications (GoA)* to refer to several use cases of application flows with similar application domains. For example, the following use-cases for a VoIP application would be classed in this paper as the GoA for *real-time audio*: one-to-one chats, many-to-many conference, one-to-many lecture.

In our recent work [2], [3], we have found that traffic characteristics for individual GoAs, such as data rate and packet size, have a great effect on the exhibited performance in WLAN operation for the different 802.11 variants. We find that the upper and lower bounds of GoA-neutral *performance envelopes* can be evaluated, empirically, across a range of flow constructions by controlled combinations of packet size and offered load. Depending on their requirements, GoA traffic profiles will lie somewhere within this envelope. For instance, to compensate for loss, real-time applications often use small packets and low data rates (with UDP). This may in turn be dependent on the use of a specific video or audio codec. Streamed multimedia flows (non-real-time), however, can compensate for loss by use of playout buffering and retransmissions, and so can operate with higher data rates and packet sizes even if there is some loss (with UDP or TCP). Bulk data transfers (e.g. file transfers) often just use as much of the available capacity as possible (using TCP, or a combination of TCP and UDP as in BitTorrent). Users currently may roam between WLAN cells without any control over the used WLAN variant on a per-client basis, and certainly have no control of use of WLAN on a per-application basis. Additionally, current hardware does not support such per-application configuration. We find that the WLAN variant(s) used in a cell, however, may impact the performance of flows within a specific GoA. We show that:

- different applications on the same client system may be better served by different WLAN variants matched to the performance requirements of different groups of applications (GoA), and so future WLAN systems may find benefit from using different WLAN variants on a per-application basis.
- our analyses of the single flow traffic profile on a testbed is useful in evaluating aggregate traffic patterns, for a group of applications (GoA), from a large network.

To investigate this behaviour for traffic in different WLAN environments, in this paper, we:

- determine performance envelopes for a wide range of traffic flows in IEEE 802.11g & 802.11n in a testbed.
- use the testbed observations to help analyse the distribution of traffic profiles from a large WLAN traffic trace, in the context terms of group of applications (GoA) .

We analyse NetFlow traces from the University of Twente comprising $\sim$5000 WLAN users of IEEE 802.11n/g radio cells. We empirically evaluate a range of GoA characteristics. We then compare these to a performance envelope, based on the operation of a single client, generated on an IEEE 802.11n/g testbed. While we have constrained ourselves to 802.11g and 802.11n for practical purposes (e.g. the network configuration of the University of Twente deployment), our methodology for generating the performance envelopes has been applied to other IEEE 802.11 variants in our previous work [2], [3]. We consider GoAs from real-time audio to bulk data transfers.

Our work is an initial comparison of use case (GoA) specific traffic patterns, extracted from a campus network, with performance envelopes generated from a WLAN testbed. The traffic patterns may vary with applications and protocols evolving, we however see a trend that multi media GoAs prefer network conditions as provided by legacy standards.

The remainder of this paper is as follows: In Section I we overview some of the related research achievements. After that in Section II we explain our methodology. Following, we show our experimental findings and the extracted traffic characteristics in Section III and provide concluding remarks and an outline to future work in Section IV.

I. Related Work

Overall, no existing studies examines the performance of specific IEEE 802.11 variants in the context of their suitability for application-specific flows.

Henderson *et. al.* [4] provide a rigorous analysis of trends of the usage in the 802.11b wireless campus network of the Dartmouth College in the US. Overall they define and consider the following categories of applications: Bulk, Database, Interactive (e.g. IRC, AIM), Mail, P2P (e.g. Gnutella), Services (e.g. X11, DNS), filesystem (e.g. SMB/CIFS), Streaming (e.g. RealAudio), VoiP (e.g. Cisco CallManager), WWW, Other (all named ports that do not fit into the other categories), Unknown (all unnamed and unidentified ports). Similarly, we also have groups of applications (GoA) which we analyse.

Fiehe *et. al.* [5] examine performance in 802.11n experimentally, considering signal strength attenuations. They do not evaluate practical upper and lower bounds of operation as determined by the flow characteristics. A similar discussion applies to Shrivastava *et al* [6], in which the authors experimentally evaluate the impact on performance of 802.11n features like MIMO, channel bonding and frame aggregation. They also consider a scenario in which the presence of a neighbouring 802.11g cell causes interference, in a specific office environment and configuration. In [7]–[9] the authors report on empirical measurements of performance in IEEE 802.11 networks. The analyses focuses on coverage, RSSI and interference. The authors provide measurements of application-specific performance using *ping* to determine loss, and file transfers to determine throughput, but they do not consider upper and lower bounds of what is practically achievable. Suong *et. al.* [10] use model based analyses to conclude that many small packets will result in an increased probability of collision (as we do).

As part of our study we also dealt with application/traffic classification – a commonly used technique, evolving over time from packet-based [11], [12] to flow-based. So, a number of studies that identify application types and assign them into specific categories/classes are popular for network management practises and academic studies [13].

Karagiannis *et. al.* proposed a method that is based on the transport layer characteristics and therefore does not rely on user payload inspection [14]. This method makes use of two heuristics: one examines the source-destination IP pairs that use both TCP and UDP to transfer data, while the second monitors connection patterns of transport endpoints. The growing applicability of clustering and machine learning methodologies pushed the research community to consider their application for the purposes of traffic classification [15], [16] as well as intrusion and anomaly detections [17], [18].

II. Metrics and Approach

A. Overview

We compare testbed results for a performance envelope, with analyses of NetFlow data from a campus network. We use the same testbed and harness as already described in [2], [3] for our WLAN performance measurements. We extract traffic profiles from NetFlow traces comprising a user base of $\sim$5000 users of 802.11n/g wireless networks operating at 2.4GHz at the University of Twente. The NetFlow data is processed to extract specific traffic profiles (specified by data rate of the offered load and packet size). The range of the values for data rate and packet size are manually evaluated for the individual groups of applications (GoA).

For our testbed, we assume that most users do not have the expertise to fine-tune their equipment and that most deployed systems are used in 'out-of-the-box' configurations (no performance tuning). Specifically, our constraints are:

- *Standard WLAN configuration.* We used only standard, un-tuned WLAN setups. While many WLAN NIC drivers and access points (AP) do permit various controls of the hardware, this is not easily accessible or comprehensible for modification by most users.
- *Packet flow behaviour.* To measure application-specific performance (throughput and loss) we use a range of UDP flows specified by packet rate and packet size to evaluate the upper and lower performance bounds that define our performance envelope.

B. Traffic Profile extraction from NetFlow Traces

We extract traffic profiles of both real-time and streamed audio and video applications as well as bulk data transfers from NetFlow formatted traces from the University of Twente.

1) NetFlow: NetFlow services provide network administrators with access to IP flow information from their data networks. Network elements (routers and switches) gather flow data and export it to collectors. The collected data provides fine-grained metering for highly flexible and detailed resource usage accounting. Generally, a flow is a set of packets that share common properties. In [19], a flow is defined as a unidirectional stream of packets between a given source and destination, both specified by a network-layer IP address and a transport-layer source and destination port numbers. As such, a NetFlow v5 flow is identified by a combination of seven connection specific fields (Source IP address & Port number; Destination IP address & Port number; IP protocol type – e.g. TCP or UDP; Type of service byte; Input logical interface)

In addition, a flow contains other accounting fields which may differ slightly depending on the NetFlow version record format. For instance, suppose that an SSH connection is established from a client on host 12.14.2.3 port 1234 to a server on host 13.18.5.6 port 22, and that the traffic passes through a router that has NetFlow processing enabled. The initial packet from the client to the server causes the router to create a flow entry for {`TCP, 12.14.2.3, 1234, 13.18.5.6, 22`}. The response from the server to the client causes the router to create a related flow {`TCP, 13.18.5.6, 22, 12.14.2.3, 1234`}. Data (packet sizes, number of packets per second, etc.) from subsequent traffic will be aggregated in these two flow records until one of the terminating conditions for the flow are met.

2) Trace Description: The NetFlow data is captured at the central router of the University of Twente in the Netherlands. This data set represents daily Internet audience of ∼5000 users. Most of these users have static IP addresses, while the others have them dynamically assigned. A certain, known, fraction of traffic originates in Wireless LANs of the university facilities, which operates with 802.11n/g at 2.4 GHz. For our work we used traces of weeks 35-36 in 2010 and weeks 8-9 in 2011. We compared the statistics for this randomly selected time period with other weeks and did not find significant differences. The longest period for which the selected set of flow records can be considered representative is seven weeks.

3) Traffic Profile Extraction: Since the identified GoAs represent a general pool of applications and protocols, we have used a simple approach for their differentiation. We define five GoAs: real-time audio & video, streamed i.e. non-real-time audio & video, and bulk data transfers. In order to assign a NetFlow record into one of these use cases we first carried out a manual investigation (using tools such as `tcpdump` and `nfdump` to identify traffic parameters (packet size and data rate) for each class. For instance, in case of real-time video, we initialised Skype and Gtalk Video sessions with different endpoints and monitored the statistics of the resulting packets and flow records. This was used to identify upper and lower bounds of expected packet sizes (bytes) and data rates (Mbps) for the real-time video use case. A similar procedure was carried out for the other four GoAs. Real-time audio streams were based on the same applications as above, and were tested with multiple endpoints. Streaming video applications included: YouTube streams of different quality (360p, 480p and 720p); Vimeo streams with 'HD-off' and 'HD-on', as well as streams from Dailymotion. For streamed audio we considered a number of on–line music players, like Last.fm and Grooveshark, as well as several on-line radio stations and browser radio plugins. To establish the parameters for bulk GoA, services like Megaupload and Dropbox in combination with P2P and torrent clients were used to measure characteristics of downloads of files of various sizes.

Based on these observations, we formed a set of five filter rules that were fed into `nfdump` along with the traffic trace. The produced output consisted of five NetFlow traces that included records associated with each of the GoAs.

Next, we used a combination of `nfdump` and `bash` scripts in order to extract the statistics for each traffic class, producing distributions of packet size values and data rate for each individual traffic class.

C. Testbed

We have experimentally measured performance in our WLAN testbed. We generated packet flows of offered loads with various packets sizes, and measured end-to-end performance during the packet transmission. Our testbed (Figure 1) consisted of a single client host, a host running a wireless access-point (AP) and experimental control units for monitoring the WLAN environment, providing storage for measurement data, `ntp` services and system configuration. The WLAN hosts were setup in a teaching lab in the University of St. Andrews with a distance of ∼24±0.5 m between the 2 dBi antennas. We have conducted our measurements using 11g and 11n at 2.4GHz to present values representative of those found in the traces.

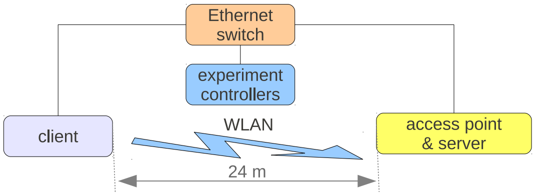

Fig. 1. Schematic of test-bed showing physical connectivity. The testbed was configured separately for 802.11g and 802.11n (20 MHz – default - channel width), both with transmission (TX) powers of 17 dBm. The experiment controller uses Ethernet for control messages and shared file-system access. The distance between the client and access point antennas is 24±0.5 m. Data packets generated by `iperf` were transferred across the WLAN link.

We tested 802.11n and 802.11g at 17 dBm (50 mW, a high but typical indoor RF power – chosen to avoid measurements being biased by poor RF conditions, and representing a best-case scenario), and with a 20 MHz (default) channel width for 11n. This means that all our experimental workloads in Table I are executed 2 times, once with each of these combinations. Our WLAN card uses the popular Atheros chipset. All machines used Ubuntu 10.04 a minimal server distribution (no desktop service daemons or GUI overhead), with the default kernel 2.6.32-24-generic-pae, and updated WLAN modules (compat-wireless-2011-05-02), which will soon be part of the standard distribution. We have used hostapd (v. 0.6.9) as AP and to avoid overhead and bias due to link encryption and security mechanisms we disabled encryption and security. To prevent experiments being disturbed by other users, our WLAN cell did not broadcast the SSID in the beacon interval.

D. Experiments

Packet generation and performance measurement for UDP traffic was conducted using `iperf` for which the AP was used as the server. A wrapper script at the client executed `iperf` and extracted throughput and loss for individual UDP flows using `iperf` *server reports*. The specific packet sizes and bit rates of the UDP workload are given in Table I, for which we choose a range into which most applications fall. We use UDP as it is popular for Voice and Video over IP (VoIP and ViIP) applications and because it allows better control of application-specific offered load bit rates compared to TCP, which is modulated by its congestion control behaviour.

TABLE I
UDP WORKLOAD.

Packet size	64; 1460 bytes
Bit rate of the offered load	10; 50; 100; 500 Kbps 1; 5; 10; 15; 20; 25; 30 Mbps

Combination of packet size and bit-rate produces 22 tuples; 5 measurements for each combination; 11g and 11n with 20MHz (default) channel width; each flow had a duration of 120 seconds, giving over 7 hours of measurements. We choose to uniformly test up to a max. of 30Mbps for offered load for comparison reasons – this is just above the operational limit of 802.11g.

III. Results and Discussion

We present the characterisation of GoAs from the traffic profiles extracted from our traces by showing the distribution of packet sizes and data rate for each individual use case i.e. for each individual GoA. We compare that with the upper and lower bounds of performance, that means, with our *performance envelope*, measured in our WLAN testbed operating separately with IEEE 802.11n and IEEE 802.11g. In addition to throughput, we also show measurements for loss, as that is often ignored in such investigations. We can summarise that our analyses allows the identification of GoAs in which the degree of variation in the extracted flow characteristics is of such magnitude that users will exhibit, performance changes when roaming between WLAN cells of different standards.

A. Use Case Specific Trace Analysis

Packet size distribution is, for each GoA, less skewed than the corresponding data rate distribution. The latter's skewness increases with the average packet size of the specific GoA. In Table II we summarise packet size and data rate distribution by presenting the mean and standard deviation for traffic for a specific GoA. Because of the skewness of the data rate distributions, we also show the mode value for the data rate distributions. We see distinct traffic profiles for each GoA. To illustrate this in more detail we show the cumulative frequency distributions for all use case specific traffic profiles and flow parameters in Figures 4–8. For our analysis we have used the trace as described in section II-B2, i.e. the data is representative of consecutive seven weeks in our trace. Distributions of traffic profiles for each GoA are skewed and have a coefficient varying from ≈ 0.4 to ≈ 3 (using Pearson's skewness definition: $3.(mean - mode)/std$). This is expected, since each GoA is biased towards a certain type of flow type. We observe that the bulk class exhibits a large value for the standard deviation for a given data rate, which is exhibited as a heavy-tail in the CFDs.

TABLE II
Throughput and packet size distribution characteristics.

	packet size [byte]		data rate [Kbps]		
GoA (use case)	avg	std	mode	avg	std
real-time audio	147	42	43	70	26
real-time video	570	96	528	588	186
streamed audio	854	72	86	171	107
streamed video	1079	159	777	1562	1118
bulk data transfer	1434	42	219	688	3429

Average data rate and packet size of use cases and classes of applications, captured during one week in 2011, comprising about 1000 unique MAC addresses – please see Section II-B2 for the trace description.

B. Flow Characteristic Dependent Performance Envelopes

Figures 2 and 3 show the measured average throughput and loss in a 802.11g and 802.11n WLANs operating at 2.4GHz, with 20MHz channel width. This shows that at a given data rate of the offered load (i.e. the application's data rate) throughput is determined by the packet size. This is a result of the MAC layer overhead (see [2] for more details). Additionally we also see that loss increases for smaller packets in 11n. In general we can say – based on these measurements and our past experience – that a throughput gain comes at the cost of increased loss. We illustrate that by including loss plots for 802.11g and 11n in Figure 2 and 3. They represent the upper and lower bounds of throughput and loss in the individual WLAN standards. Performance of real applications will be somewhere within these *performance envelopes*.

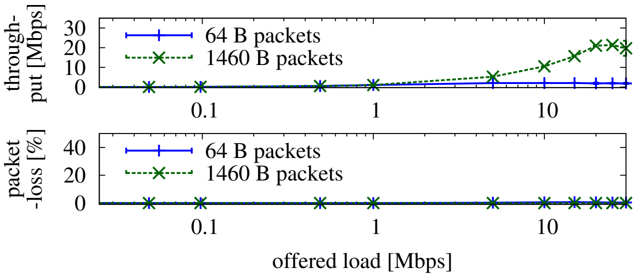

Fig. 2. IEEE 802.11g with 17dBm TX power. Each point represents 5 flows with the same data rate, packet size and of 120 sec duration.

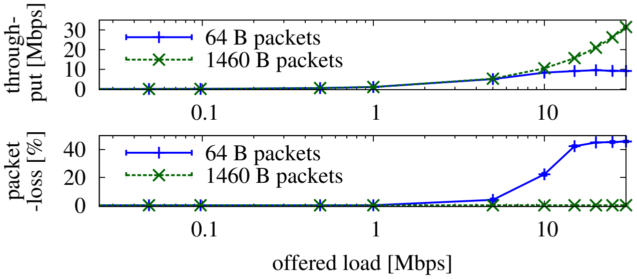

Fig. 3. IEEE 802.11n with 20MHz chan. and 17dBm TX power. Each point represents 5 flows with the same data rate, packet size and of 120 sec duration.

C. Comparison of Performance Envelopes and Trace Analysis

When putting the performance envelopes in Figure 2 and 3 in context with our extracted traffic profiles (Table II) we see that streamed video traffic operates (due to the high standard deviation of packet sizes) in a mode where loss can more likely occur due to the applications flow behaviour if being used in 11n but not with 11g. Even though bulk data transfer traffic can have a high data rate it may not suffer from too much loss as the packet size does not vary as much as it does for streamed video. In our traffic analysis we see that bulk data transfer exhibits data rates below 11g's operational limit which does not result in a difference in performance in our envelope of different WLAN flavours. However, in e.g. Intranet use cases (not isolated in our analyses) local caches may result in higher data rates and lower end-to-end path loss, which will, potentially, result in superior performance in 11n than in 11g. Real time applications, however operate with such flow characteristics that no performance difference due to tested standards is to be expected when e.g. roaming between 11g and 11n. That is, from our performance envelopes, we see that at loads below ∼1 Mbps, there is little difference in throughput and loss across 11n and 11g, but above ∼1 Mbps differences start to appear, and they will impact different GoAs differently.

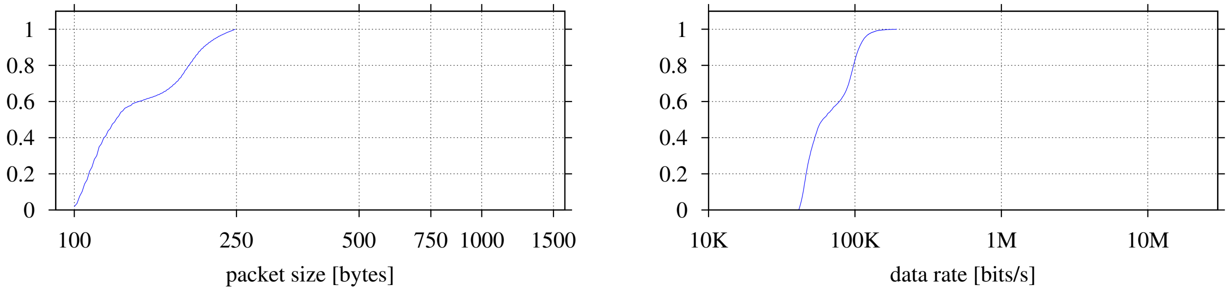

Fig. 4. Cumulative Frequency Distributions of Traffic Flow Characteristics extracted for real-time audio.

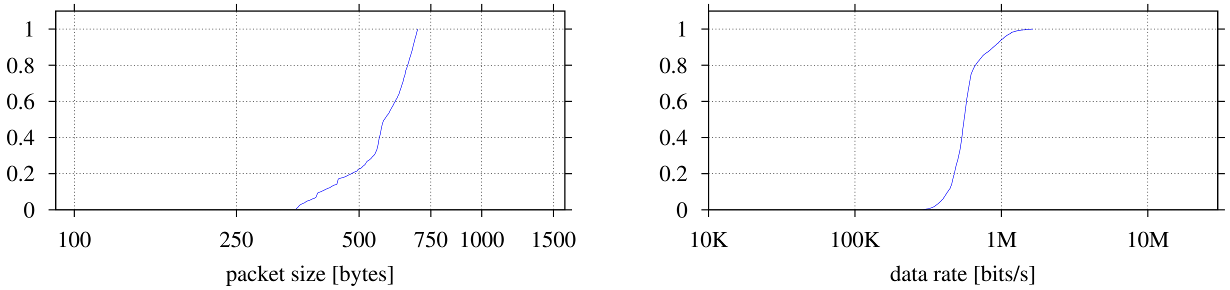

Fig. 5. Cumulative Frequency Distributions of Traffic Flow Characteristics extracted for real-time video.

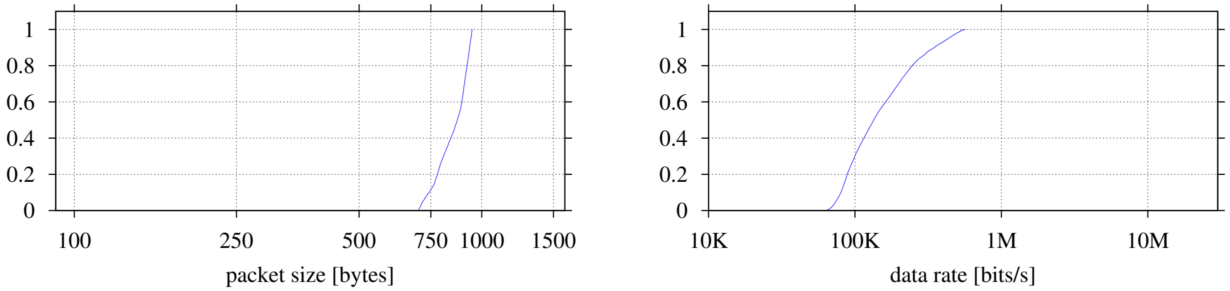

Fig. 6. Cumulative Frequency Distributions of Traffic Flow Characteristics extracted for streamed audio.

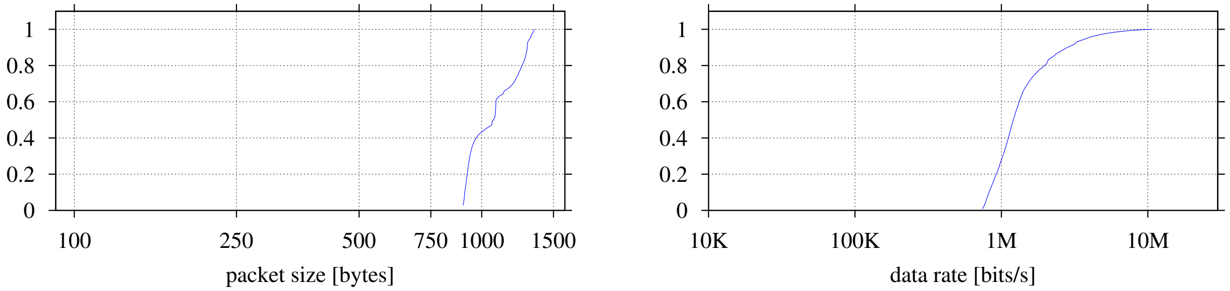

Fig. 7. Cumulative Frequency Distributions of Traffic Flow Characteristics extracted for streamed video.

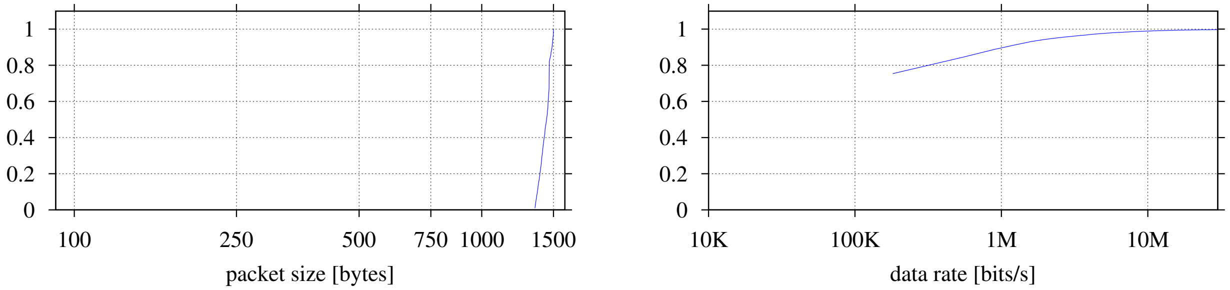

Fig. 8. Cumulative Frequency Distributions of Traffic Flow Characteristics extracted for bulk data traffic.

D. Practical implications of our results

We see from our measurements that with higher offered load and larger packet sizes (Figure 2 and 3), although higher throughput is achieved, there is comparatively higher loss for 802.11n, but not for 802.11g. Such loss is not suitable for real-time audio and video. Indeed, our characterisation of the traffic (Figures 4–8) show that real-time audio has smaller packets sizes – less than 260 bytes – and lower data rates – less than ∼160Kbps. For real-time audio, the relevant figures are packet sizes of less than 700 bytes and data rates of up to ∼1.5Mbps with the 90th-percentile at ∼880Kbps. For streamed audio and video, we see higher data rates and higher packet sizes being used, as these applications can tolerate loss. In 11g, we see from our testbed that loss is not such a factor at equivalent data rates of 11n (Figure 2 and 3). So, it is possible that a mixed 802.11 environment might exist in the future. The lower bit-rate, real-time applications would use 802.11g to exploit low loss, and the higher bit-rate streamed applications would use 802.11n to use the higher data rates. However, the client systems would need to support simultaneous use of 11g and 11n, something that is not widely supported today (client interfaces operate in either 11g or 11n, by access point interfaces can use both). Another factor would be if the loss characteristics of 11n could be improved at higher data rates.

IV. Conclusions and Future Work

From our previous work on WLAN performance [2], [3], we find that application specific data rates and packet sizes impact traffic. Here, we have characterised traffic from a campus wireless network by extracting traffic profiles of five groups of applications (GoA): real time audio, real-time video, streamed audio, streamed video, and bulk data transfer applications. The flow characteristics of the GoA profiles have then been put into context by examining performance measurements conducted in our local testbed. We have established *performance envelopes* which show the upper and lower bounds of performance in 11g and 11n. (restricted to 11g and 11n to match the analysed traces from the campus network). We see that specific GoA profiles have a greater probability of suffering from loss in specific WLAN environments. Especially multimedia GoAs seem to benefit from being used with legacy 802.11 variants. Comparing the distributions of traffic characteristics of the GoAs to our performance envelope for 11g and 11n, we find that real-time traffic may be better suited to 11g (lower loss), not benefiting from the higher rates of 11n (with higher loss). Non-real-time and bulk data, however, can make better use of 11n.

Our analyses shows that different GoAs may be better served by different 802.11 variants. So, future deployments may wish to exploit this through a parallel-mode deployment of 11g and 11n, allocating these to different *applications*, something which is not done today. Alternatively, applications could be enabled to adapt their flow characteristics to match 802.11 variant capability.

Future work includes measurements with multiple clients and dual-mode operation and use of different end-system platforms, testing various other standards as well as an extended trace analysis and comparisons with other traffic studies.

Acknowledgements

We are grateful to our colleagues from Twente University for access to NetFlow data. The work at St Andrews has been supported mainly by the EPSRC-funded project IU-ATC.

References

[1] V. Perelman, N. Melnikov, and J. Schönwälder, “Flow signatures of popular applications,” in *12th IFIP/IEEE Intl. Symposium on Integrated Network Management (IM) 2011*, May 2010.

[2] M. Tauber, S. N. Bhatti, and Y. Yu, “Application Level Energy and Performance Measurements in a Wireless LAN,” in *Proc. GreenCom2011: IEEE/ACM Intl. Conference on Green Computing and Communications*.

[3] M. Tauber and S. N. Bhatti and Y. Yu, “Towards Energy-Awareness in Managing Wireless LAN Applications,” in *Proc. NOMS 2012: IEEE/IFIP Network Operations and Management Symposium*, 2012.

[4] T. Henderson, D. Kotz, and I. Abyzov, “The changing usage of a mature campus-wide wireless network,” *Computer Networks*, vol. 52, no. 14, pp. 2690 – 2712, 2008.

[5] S. Fiehe, J. Riihijärvi, and P. Mähönen, “Experimental Study on Performance of IEEE 802.11n and Impact of Interferers on the 2.4 GHz ISM Band,” in *Proc. of the 6th Intl. Wireless Communications and Mobile Computing Conference*, ser. IWCMC ’10. New York, NY, USA: ACM, 2010, pp. 47–51.

[6] V. Shrivastava, S. Rayanchu, J. Yoonj, and S. Banerjee, “802.11n under the microscope,” in *Proc. of the 8th ACM SIGCOMM conference on Internet measurement*, ser. IMC ’08. New York, NY, USA: ACM, 2008, pp. 105–110.

[7] D. J. Leith and D. Malone, “Field Measurements of 802.11 Collision, Noise and Hidden-Node Loss Rates,” in *WiOpt’10: Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Avignon, France, 2010, pp. 521–526.

[8] L. Verma, D. Sim, C.-Y. H. Yang, K. Cho, and S. S. Lee, “On Measurement Campaign with IEEE 802.11n Devices,” in *Proc. IEEE ICOIN’10*, 2010.

[9] S. Sendra, P. Fernandez, C. Turro, and J. Lloret, “IEEE 802.11a/b/g/n Indoor Coverage and Performance Comparison,” in *6th Intl. Conference on Wireless and Mobile Communications (ICWMC)*, 2010.

[10] S. H. Nguyen, H. L. Vu, and L. L. H. Andrew, “Packet Size Variability Affects Collisions and Energy Efficiency in WLANs,” in *2010 IEEE Wireless Communications and Networking Conference (WCNC 2008).*

[11] L. Bernaille, R. Teixeira, I. Akodkenou, A. Soule, and K. Salamatian, “Traffic Classification on the Fly,” *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 2, pp. 23–26, 2006.

[12] S. Sen, O. Spatscheck, and D. Wang, “Accurate, Scalable In-network Identification of P2P Traffic Using Application Signatures,” in *WWW ’04: Proc. of the 13th international conference on World Wide Web*. New York, NY, USA: ACM, 2004, pp. 512–521.

[13] A. Moore and K. Papagiannaki, “Toward the accurate identification of network applications,” in *Proc. PAM 2005*, 2005, pp. 41–54.

[14] T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy, “Transport layer identification of P2P traffic,” in *IMC ’04: Proc. of the 4th ACM SIGCOMM conference on Internet measurement*. New York, NY, USA: ACM, 2004, pp. 121–134.

[15] M. Perényi, T. D. Dang, A. Gefferth, and S. Molnár, “Identification and analysis of peer-to-peer traffic,” *JCM*, vol. 1, no. 7, pp. 36–46, 2006.

[16] B. Park, Y. J. Won, M.-J. Choi, M.-S. Kim, and J. W. Hong, “Empirical analysis of application-level traffic classification using supervised machine learning,” in *APNOMS ’08: Proc. of the 11th Asia-Pacific Symposium on Network Operations and Management*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 474–477.

[17] A. Lakhina, M. Crovella, and C. Diot, “Mining anomalies using traffic feature distributions,” *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 4, pp. 217–228, 2005.

[18] Y. Gao, Z. Li, and Y. Chen, “A doS resilient flow-level intrusion detection approach for high-speed networks,” in *ICDCS ’06: Proc. of the 26th IEEE International Conference on Distributed Computing Systems*. Washington, DC, USA: IEEE Computer Society, 2006, p. 39.

[19] B. Claise, “Cisco Systems NetFlow Services Export Version 9,” Cisco Systems, RFC 3954, Oct. 2004.