Treatment-Based Traffic Signatures

Mark Claypool, Robert Kinicki, Craig Wills
Computer Science Department at Worcester Polytechnic Institute
100 Institute Road, Worcester, MA, 01609, USA
{claypool,rek,cew}@cs.wpi.edu

As depicted in Figure 1, today's wireless local area networks (WLANs) concurrently support a plethora of user applications with diverse connectivity and Quality of Service (QoS) requirements. Wireless traffic flows contending for resources can degrade performance for one or more of the wireless applications. Delay sensitive applications may face high latency in the presence of competing wireless traffic that saturates the WLAN. Streaming applications that need an available bandwidth estimate to select the best encoded media for good performance face challenges doing bandwidth estimation over wireless networks. Requiring smooth data delivery, VoIP suffers from delay jitter during WLAN congestion.

One solution to these performance issues is a 'smart' Access Point (AP) that automatically improves performance in an interoperable, easy-to-use fashion that a user expects. However, prior to designing such an AP, techniques that identify and classify applications based on wireless network traffic characteristics are needed to enable the AP to respond accordingly to diverse QoS demands. A major component of this research approach is developing representative traffic signatures to facilitate classifying WLAN traffic for subsequent network treatments.

Deriving traffic signatures that move beyond simple, error-prone, port-based heuristics is a topic of recent study. Roughan et al. [4] define signatures based on average packet sizes and flow duration. BLINC [2] classifies hosts rather than individual flows and does so at the social, functional and application level. Other work [1, 3, 5, 6] focuses on accurate identification of the applications themselves.

Rather than classify applications, our approach concentrates on the *nature of traffic* due to specific applications and devices. The distinction is two-fold: 1) Different applications with the same QoS requirements should receive equivalent network treatments. Separate classification amongst such applications is unnecessary and may be harmful when slow or intrusive efforts are taken to provide the distinctions; and 2) Not all instances of a particular application yield the same signature, nor is that needed. For example, consider two instances of a Web application, where the first retrieves a large image for display using relatively large packets over a long duration and the second retrieves an index page with many small embedded images consisting of small amounts of data over a short duration. Both are instances of the same application, but classifying them as having the same application signature is difficult and not useful for meeting their network QoS requirements.

Figure 1: The Diversity of Internet Applications

By characterizing flow signatures in terms of the nature of the traffic, our research objective is to develop a better and more effective matching strategy between the resultant signatures and the potential treatments that need to be applied to WLAN traffic flows. Thus, we propose traffic characterization based on three classifiers:

1. Nature of Reverse Traffic. Is the traffic uni- or bi-directional and is the traffic that flows in the reverse direction response-based? Response-based, which includes ACK-based, traffic means that all packets flowing in one direction will directly or indirectly (e.g. delayed TCP ACKs) cause traffic flow in the reverse direction. By definition, applications built on TCP are response-based with its ACKs, but UDP-based applications such as DNS or any RPC requests are also response-based.

2. Packet Size Tendency. Do packets in the flow tend to be full, meaning they are the maximum size sup-

ported by the network, or do they tend to be non-full, meaning they are less than the maximum size? This classifier is similar to classification based on the average packet size [4], but examines the ratio of full to non-full packets, where bulk-data transfers are likely dominated by full packets and response-oriented applications tend to send smaller, non-full packets.

3. Transmission Spacing of Packets. Does the pattern indicate packets are transmitted in bursts on an 'as-available' basis or does the pattern indicate a paced spacing in transmission? While the former often indicates both throughput- and response-oriented applications, the latter pattern is indicative of applications that require a steady data rate to limit jitter. Congestion and queuing are potential impediments to correct determination of this classifier, but we believe the long-term nature of these applications makes classification feasible.

These three classifiers form the basis of our classification and integrate with the treatments that can be applied to each traffic signature. The classifiers form axes in a classification space for both applications and potential treatments shown in the cube of Figure 2. Focusing first on representative applications, which are overlaid as shaded areas on the face of the cube, most instances of an application are expected to consistently be classified in the same manner. For example, interactive applications, such as Telnet or DNS, exhibit response-based, non-full packet traffic that is sent as it is available. Instances of other applications, such as Web or Games, span multiple portions of the classification cube. Rather than this 'multiple classification' being a problem with the approach, it is instead a feature as it properly indicates that not all instances of an application need be treated in the same manner. The key point in showing the application classification in Figure 2 is not so much to classify the applications as it is to indicate the *expected range* of instances of an application.

The network treatments, and their range in the classification, are shown in Figure 2 with the ALLCAP font indicating the treatment and the }'s showing their range. These treatments represent general approaches that can be applied to packets within a classified traffic flow that are intended to address user-level QoS concerns. They are also attractive because they map directly to techniques that can apply to application traffic flowing through it. Four types of treatments are shown:

1. Drop packets. Packets within non-response-based flows can be dropped if needed as these applications are more robust to lost data. Drops done to response-based flows incur retransmissions, thus reducing the benefits of dropping packets.

2. Delay packets. Packets in throughput-based flows indicated by response-based, mostly full packets, can be delayed without greatly impacting application performance.

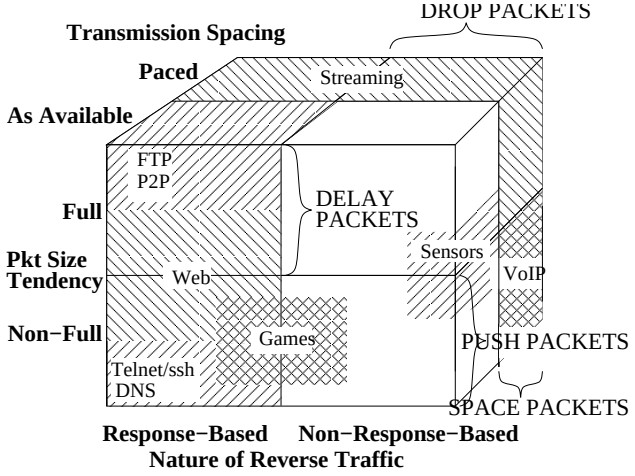

Figure 2: Possible Treatments Based on Classification

3. Push packets. Packets in flows from interactive applications, identified by non-full packets, can be pushed ahead to reduce latency. This treatment is appropriate for both response- or non-response-based flows.

4. Space packets. Packets in flows that need a consistent rate with little jitter, identified by the spacing of packets from the sender, can be pushed or delayed to maintained a consistent rate and/or delay as the network load varies. This treatment is appropriate for streaming video (VoD) and audio (VoIP) applications that need bandwidth estimation and low delay jitter.

In another component of our work these four treatments are mapped to specific techniques at the wireless, network and transport network layers to produce the desired effect on traffic flows. In summary, we believe our classification demonstrates the desirability of developing a signature approach that is driven by how it will be used.

[1] P. Haffner, S. Sen, O. Spatscheck, and D. Wang. ACAS: Automated Construction of Application Signatures. In *ACM SIGCOMM Workshop on Mining Network Data*, New York, NY, USA, 2005.

[2] T. Karagiannis, K. Papagiannaki, and M. Faloutsos. BLINC: Multilevel Traffic Classification in the Dark. In *Proceedings of ACM SIGCOMM*, New York, NY, USA, 2005.

[3] A. Moore and K. Papagiannaki. Toward the Accurate Identification of Network Applications. In *Passive and Active Measurement Workshop (PAM)*, Boston, MA, USA, Mar/Apr 2005.

[4] M. Roughan, S. Sen, O. Spatscheck, and N. Duffield. Class-of-service Mapping for QoS: a Statistical Signature-based Approach to IP Traffic Classification. In *ACM SIGCOMM IMC*, New York, NY, USA, 2004.

[5] S. Sen, O. Spatscheck, and D. Wang. Accurate, Scalable In-Network Identification of P2P Traffic Using Application Signatures. In *WWW*, New York, NY, USA, 2004.

[6] A. Soule, K. Salamatia, N. Taft, R. Emilion, and K. Papagiannaki. Flow Classification by Histograms: or How to Go on Safari in the Internet. In *SIGMETRICS/Performance*, New York, NY, USA, 2004.