Data Mining - A Captured Wired Traffic Approach

Sasan Adibi
ECE Department, University of Waterloo
sadibi@uwaterloo.ca

Abstract

Data mining is involved in procedures by which patterns are extracted from data. This process has become increasingly important to map data patterns to useful information that can be used to predict future traffic analyses. Other areas were data mining can be used include: fraud detection, marketing, congestion control, and network expansion consideration.

Data mining involves capturing and gathering random data from the flow of information passing through a certain trunk or node and from a statistical point of view, it can only be meaningful if enough samples are gathered. The results achieved from a proper deployment of a data mining method provide valuable insights to how busy a node is, the average one-way and end-to-end delays, and the average size of the packets..

This paper contains the results from capturing live wired traffic and analyzing the statistical values.

Keywords: *Data Mining, Flow Duration, Packet Size, Confidence Interval (CI), Wireshark, Scatter Plot.*

1. Introduction

Data mining is used to categorize traffic flows based on packet information (e.g., flow duration, packet size, etc. There are numerous packet classification categories based on how packets are decoded and analyzed, including the consideration of packet size, duration, patterns of transmission and burstiness.

The management of this paper is as followed: Section II will discuss data mining parameters. Sections III is dedicated to real-time data measures, followed by conclusion, appendix and references.

2. Data mining parameters

In this section we introduce a number of network traffic parameters considered in data mining. These parameters are mostly used in the study of packet and traffic classification techniques, data mining acquisition and dissemination.

2.1. Packet size

Packet size is an important parameter in data mining. Most of the traffic flows on the Internet can be categorized into either, very large (elephant or heavy tailed) or very small (mice) packets sizes. The large packet size is usually associated with lengthier usage of a link. Basically 20% of the connections on the Internet are responsible for 80% of the traffic passing through a single node, mostly including elephant packets [1,2,3].

In regards to packet size, Zipf's law characterizes the occurrence frequency of certain packet sizes as a function of its rank in the frequency table [4]. This means that there is an imbalance in the network since 20% of the connections carry 80% of the traffic (heavy packets) and the rest of the 80% of the connections contain small packet traffic.

Traffic Engineering (TE) [5] is a concept referring to a systematic process where traffic flows are arranged into "classified" groups for simplifying the transmission throughout networks and decreasing the chance of congestions. TE is inherently capable of dealing with very large volume traffics through traffic aggregation (grouping similar traffic flows into larger traffic trunks). However one downside to TE is its low performance and inefficiency in dealing with mice flows. Another drawback of TE in regards to data mining is the fact that traffic in large and random environments (e.g., the Internet) would exhibit volatility some traffic flow specifications, such as; bandwidth and volume [5]. Fluctuations and the unsteadiness of the traffic reduce the efficiency of TE in the process of data mining.

As mentioned, in many network scenarios, traffic flows exhibit inherent bandwidth fluctuations, which creates complications in the data mining processes, leading to frequent reclassification, which reduces the data mining process performance. These fluctuations are due to the following factors [5]:

Link Exhaustion and Connection Termination: The duration of a connection can be modeled as a stochastic variable dependant on the following parameters [6,7]: The current connection (k^{th}) time duration, the current (k^{th}) connection arrival time, the protocol in use, and the client/server performance metrics (e.g., round-trip delay, client delay, server delay, etc). The effect of these parameters contributes to the variation of the flow's median time. The median time for elephant flows (heavy-hitters) increases since according to reference [8], the longer the connection duration, the higher the probability for the link to continue its connection.

Bandwidth Fluctuations: In wireless networks, bandwidth fluctuations are relatively frequent compared to those in wired networks, mostly due to RF (Radio Frequency) issues. In wired networks, bandwidth fluctuations may occur because of some other reasons, including; congestion periods and sudden increase of user demands. Other reasons behind bandwidth fluctuations in wireless networks may include: Limitations of the bandwidth availability in multi-user environments, handoff and handover between Access Points (APs), physical layer issues (e.g., reflections, refractions, multipath, etc), vulnerability to various interferences, and performance dependency to the client's (wireless user) distance to the server (AP).

Burstiness Effect: Burstiness is defined as the variation packets inter-arrival times. When the variation is high, it points to a higher value of burstiness. Multimedia data, in particular video traffic, can be affected by the traffic flow's burstiness, reflected by a number of parameters, including [8]: The temporal Auto-Covariance Function (ACF) and the Peak-to-Average Ratio (PAR). Burstiness is time sensitive and from the probability point of view, it is more probable that burstiness become an issue for a heavy-hitter connection compared to the case in a mouse flow connection.

2.1.1. Heavy hitters (elephants) versus light hitters (mice) packets: Heavy hitters can be characterized by their long durations of connections and large packet sizes. It has been noted in the literature [9,10] that there is a strong correlation between the rate of a stream, its packet sizes, and the protocol in use.

In wired traffics, from the packet size point of view, packets are usually between a few tens of bytes up to 1514 bytes. Limited by the Maximum Transmission Unit (MTU), large data files are usually broken down into various fragments before transmission. Analyzing captured real traffic, we noticed that control packets (packets containing control commands), which do not usually contain any data payloads, are less than 200 bytes and data packets with payloads are usually above 200 bytes. Heavy hitter packets, based on the data we have gathered, from packet size point of view, are packets with payloads of 300 to 1514 bytes.

Wireless-based packets, on the other hand, start from 14 bytes (e.g., ACKs "Acknowledgements", CTS "Clear-To-Send", etc., which have no data payloads), and go up to 1530 bytes, which is the limit by which, fragmentation occurs. According to our real traffic analyses, packets with over 400 bytes in lengths are labeled as heavy hitters.

2.2. Flow Duration

Duration of packet streams is another data mining parameter. Depending on the application/protocol in use, a short lived connection (a mouse flow) packet can last from a few milliseconds up to a few minutes. A long-lived connection (an elephant flow), on the other hand, can last from a few minutes up to several hours. Statistics [9,10] show that there are direct links between larger packet sizes and longer durations. Based on captured real traffic from connections containing multimedia traffic, most control packets are light connections (mice or tortoises) and other data packets are considered heavy hitters (elephants or dragonflies).

2.3. Confidence Interval (CI)

CI is a statistical parameter, which is an interval estimator [11,12,13]. Confidence intervals are used give an estimate on how reliable a sample is. For an extreme diverse sample space, such as the Internet's backbone traffic patterns, either a traffic trunk has to be monitored for an extremely long period of time (e.g., months or years) and then utilize data mining techniques on the stored data traces to find the traffics' statistical values, or use a smaller and random sample space with the aid of a confidence interval estimator. A 95% confidence interval of a population is a relatively good estimation and 99% is the optimal estimation. Bayesian and Gaussian interval estimations are examples, by which confidence intervals can be estimated.

Eq. 1 shows the CI's estimation value, which is dependent on the standard-deviation of the samples and the square-root of the number of the samples (n) in the sample space. The calculation result from the Eq. 1 will yield a 99% CI estimation value.

$$CI \approx 2.57 \times STD \div \sqrt{n} \qquad \textbf{(Eq. 1)}$$

3. Real-time data measures

In this section we analyze real captured data using Wireshark and Omnipeek Software systems. Wireshark [14] is used to sniff wired transmitted packets and Omnipeek is designed to sniff air transmitted packets (wireless).

3.1. Wired data analysis

Wireshark is a special tool designed for line sniffing. We have probed a data trunk (with prior permission) that is serving a network with tens of desktops, laptops, controllers, and access points (APs). These devices are used for running multimedia-rich and real-time, as well as non-real-time applications.

For the validity of real-traffic monitoring, Wireshark was run for a long period of time capturing live traffic passing through the network. We noticed that 4 major time periods during a working day had significant unique traffic patterns:

8 AM – 9 AM: During this time period, work-related (e.g., university-based) emails and text-based applications (e.g., remote login) are mostly being used by users. Therefore smaller packet sizes are encountered with less multimedia contents.

12 PM – 1 PM: In this time period, messenger services, multimedia applications (e.g., YouTube) and fun-related data (e.g., Yahoo, MSN, etc) are mostly used. Captured packets show mostly large packet sizes with multimedia related payloads (e.g., RTP, UDP, TCP, codecs, etc).

4 PM – 5 PM: During this period of time, more text-based and graphic-based attachments (e.g., Microsoft Word, Adobe Acrobat files, etc) are visible in the packets' payloads. Again relatively large packet sizes with less multimedia contents are observed.

12 AM – 1 AM: In this time frame most of the data transfers were conveying backup information with average size packets.

For traffic classification, 32,147 packets were randomly selected from a pool of captured data (+10 GB) from all four groups. This reflects more realistic traffic selections.

3.1.1. Data analysis: The traffic statistics are shown in Table 1. There are 32,147 packets counting for a total of 21,745,794 bytes, during a random sniffing duration of 411.820 seconds. The average number of packets per seconds is 78.059, yielding 0.422 Mbps of throughput, transferring 52,802.949 bytes per seconds on average with 676.449 bytes per average packet size.

Table 1. Average Data Statistics

Traffic Parameter	Parameter Values
Packets	32,147
Bytes	21,745,794
Sniff Duration	411.829 sec
Average packets/seconds	78.059
Average Mbit/Sec	0.422
Average bytes/sec	52,802.949
Average packet size	676.449 bytes

3.1.2. Connection durability: As mentioned earlier, heavy hitters are categorized as per packet size and as per duration length. Ongoing connections are the ones which have increased duration lengths, such as multimedia streams. Table 1 shows 32,147 packets were under observations, which are grouped for the most occurring protocols under usage. The protocols carrying data are: IPv4, TCP, and UDP. Table A.1 (Appendix A) shows the protocol hierarchy. Table A.2 shows the Ethernet packet statistics ordered by the packet sizes.

IPv4 - There are 28,166 packets out of 32,147 packets (87.62% of the total number of packets) using IPv4 and only 7 packets (0.02%) used IPv6. These 28,166 are generated by 305 individual IP addresses (end-points) on the network. Wireshark is able to classify the number of IP packets generated per node. Table A.3 shows the statistics of the nodes and the generated IP packets associated to each node, sorted from the highest number of IPv4 packets (e.g., 25,903 IPv4 packets generated by one node) showing in the top of the list down to 27 packets. This list also represents the heavy-hitters. For privacy reasons the actual end-point IP addressed are removed.

Most of the heavy-hitters on the top of lists are nodes participating in ongoing multimedia streaming over a large period of time. These nodes are not only categorized as heavy-hitters from the time period point of view, however they are usually accompanied by TCP or UDP packets with relatively large amount of payloads, which would qualify them as heavy-hitters from packet size point of view as well.

As an observation, the first item on the list has contributed to 25,903 IPv4 packets carrying 21,066,311 bytes, each packet 813 bytes on average. Another observation points to the fact that this node has been receiving more packets than sending. This is obvious from the fact that 15,513 packets were received compared to 10,390 packets that were sent. A closer look at the traces shows that this node in the receiving mode has received relatively large packets (mostly containing data payloads) and in the sending mode has sent relatively small packets (mostly containing acknowledgments).

TCP - The same scenario is true for TCP. Table A.4 shows ordered heavy hitter TCP end-points. It should be noted that besides IP address, TCP uses port numbers associated to specific applications.

UDP - Both UDP and TCP are transport protocols therefore they follow the same set of parameters. The UDP heavy hitter end-points are listed in Table A.5.

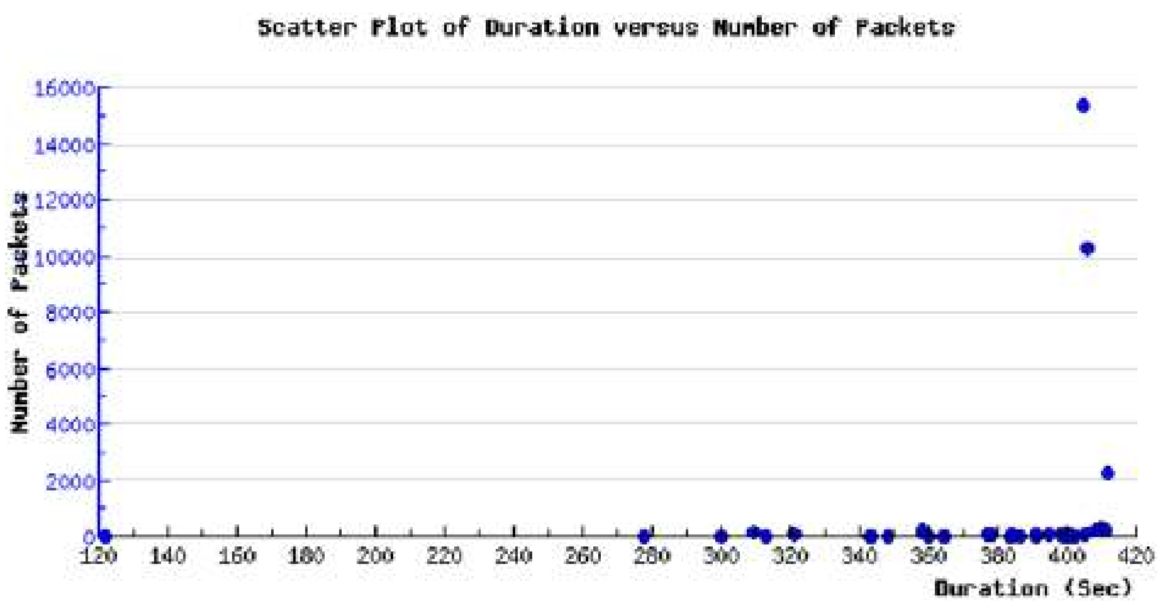

Figure 1. Scatter plot of Ethernet packets, duration vs. the number of transmitted packets

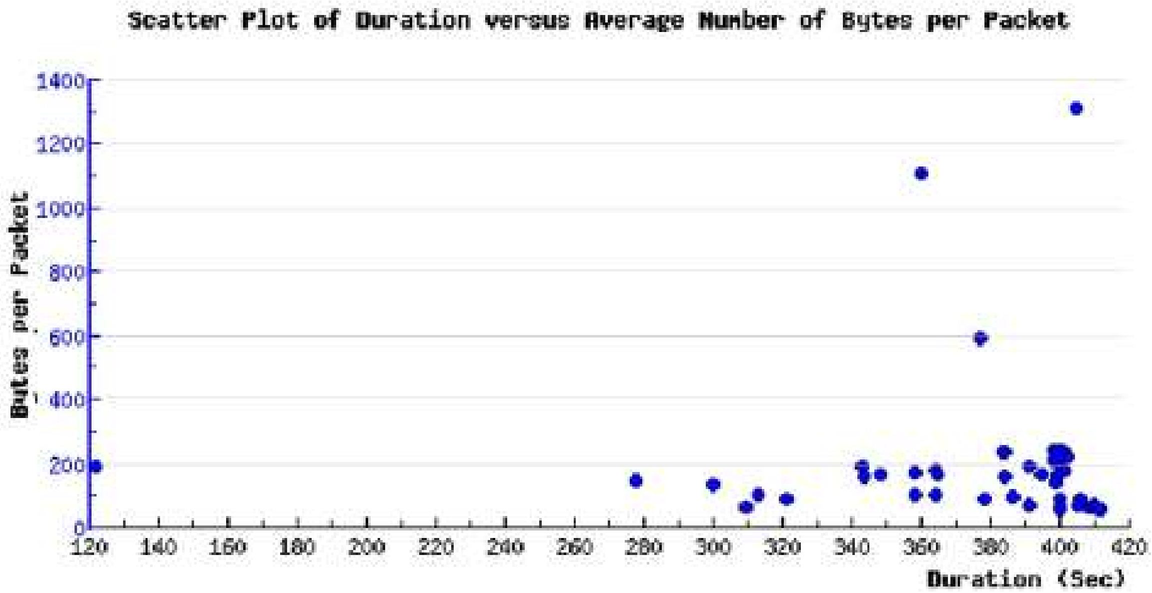

Figure 2. Scatter plot of Ethernet packets, duration vs. the number of transmitted bytes per packet

3.2. Scatter plots in data mining

Scatter plots are very useful in terms of representing the population of two and three dimensional data in respect to one another. Based on the Table A.2, the two-dimensional (duration, number of packets) scatter plot for Ethernet traffic is shown in Figure 1 [15].

The pattern in Figure 1 shows that only a few heavy hitters exist (less than 5%). The average duration time is approximately 360 seconds.

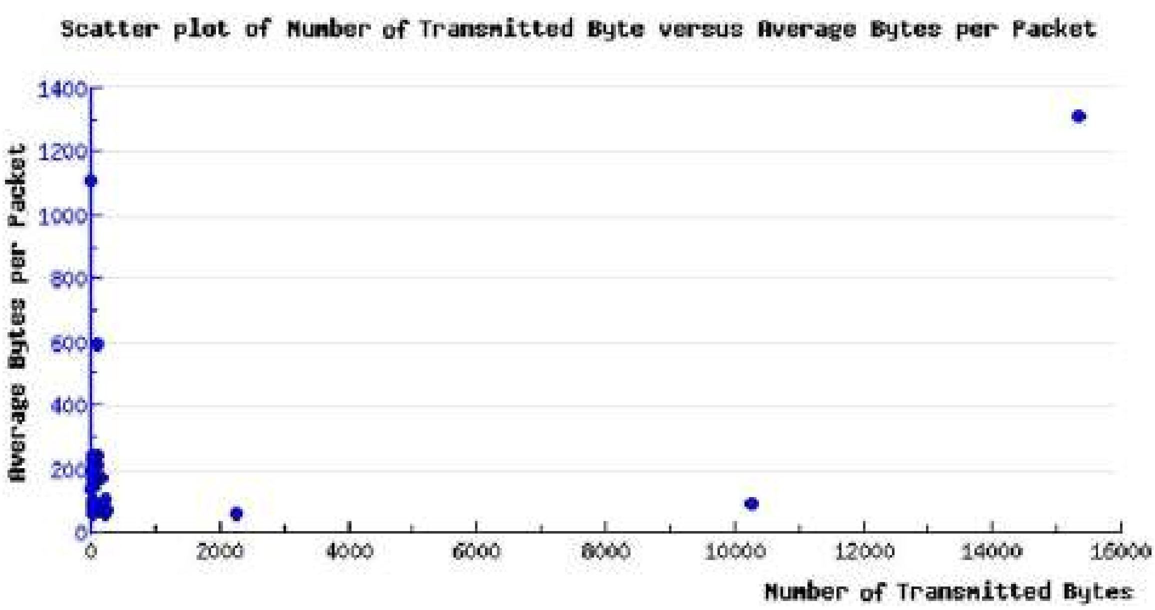

Figure 3. Scatter plot of Ethernet packets, number of transmitted bytes per packet vs. average bytes per packet

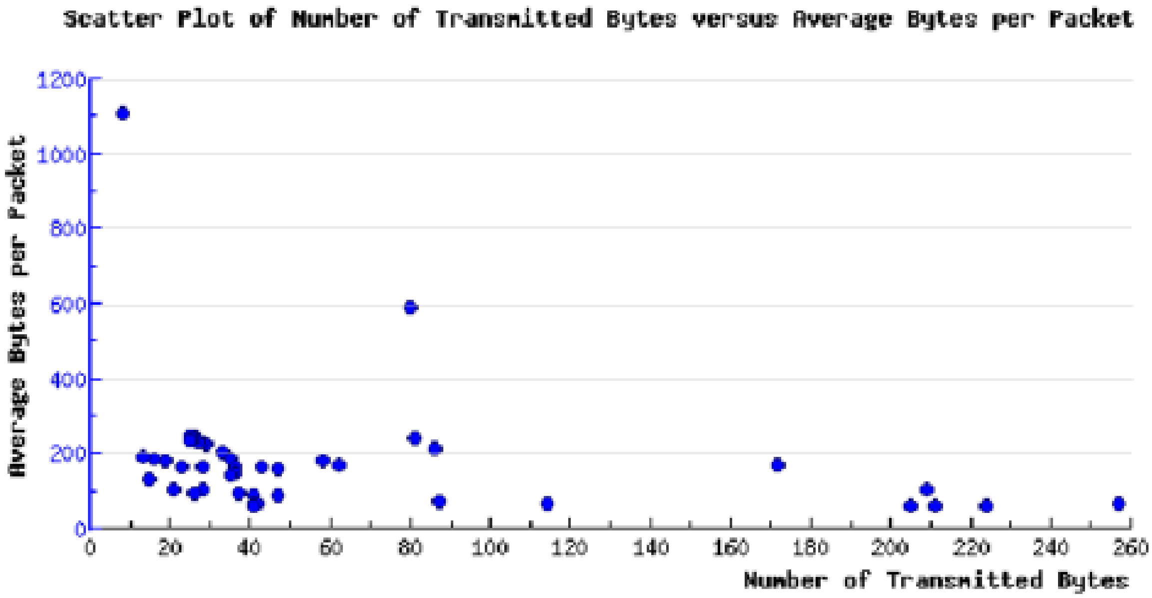

Figure 4. Expanded scatter plot of Ethernet packets, number of transmitted bytes per packet vs. average bytes per packet (based on Figure 3)

Each flow specified in Table A.2 may be associated to a different packet size. By dividing the number of packets to the number of transmitted bytes, we will get the average number of bytes per packet. For example in the first flow of Table A.2, the average number of bytes per packet is 20,079,465÷15,340 = 1,309 bytes per packet, where as in the second flow, we have: 948,450÷10,272 = 92 bytes per packets, which shows quite a large difference. The first flow may include more multimedia-rich payloads, which are often heavy in size, whereas the second flow may include more control/management frames. Figure 2 shows the scatter plot for flow-based average bytes per packet, which shows a relatively high density graph around the 340-410 seconds duration and average number of bytes between 150 to 250 bytes.

Figure 3 shows the scatter plot for the number transmitted byte versus the average bytes per packet per flow. Figures 3 and 4 are the same, except for the first three points (15,340, 1,309), (10,272, 92), and (2,262, 60). Without these three points, the rest of the scatter plot points are more expanded and visible. Figures 3 and 4 show that the density of the graph increases around the 10 to 50 transmitted bytes for the average bytes per packet of 100 to 250 bytes.

3.3. Throughput graphs

Table A.1 shows the hierarchy of protocol, as mentioned before. Ethernet frames form 99.6% of the total traffic. Cisco ISL (Inter-Switch Link) traffic form the other 0.04%.

From the 99.6% of the Ethernet traffic, 87.62% belongs to IPv4 and the rest of 12.34% belongs to other traffic types, including: Logical Link Control, Configuration Test Protocol (loopback), IPv6, and etc.

From the 87.62% of IPv4 traffic, 78.81% belongs to TCP, 8.36% belongs to UDP, and the rest of 0.45% belongs to other types of traffic.

Figures 5, 6, and 7 show the bandwidth usages for IPv4, TCP, and UDP. IPv4 and TCP graphs show high degree of correlation. This is due to the fact that the streaming data is transmitted using TCP on top of IP. Furthermore TCP is the main transport protocol used in

HTTP connections. UDP on the other hand, seems to be less correlated to IPv4, due to the fact that UDP is mainly used for controlling purposes.

Figure 8 shows the comparative bandwidth usage between HTTP, UDP, and ARP messages. HTTP uses TCP and ARP is a layer II protocol (MAC layer service).

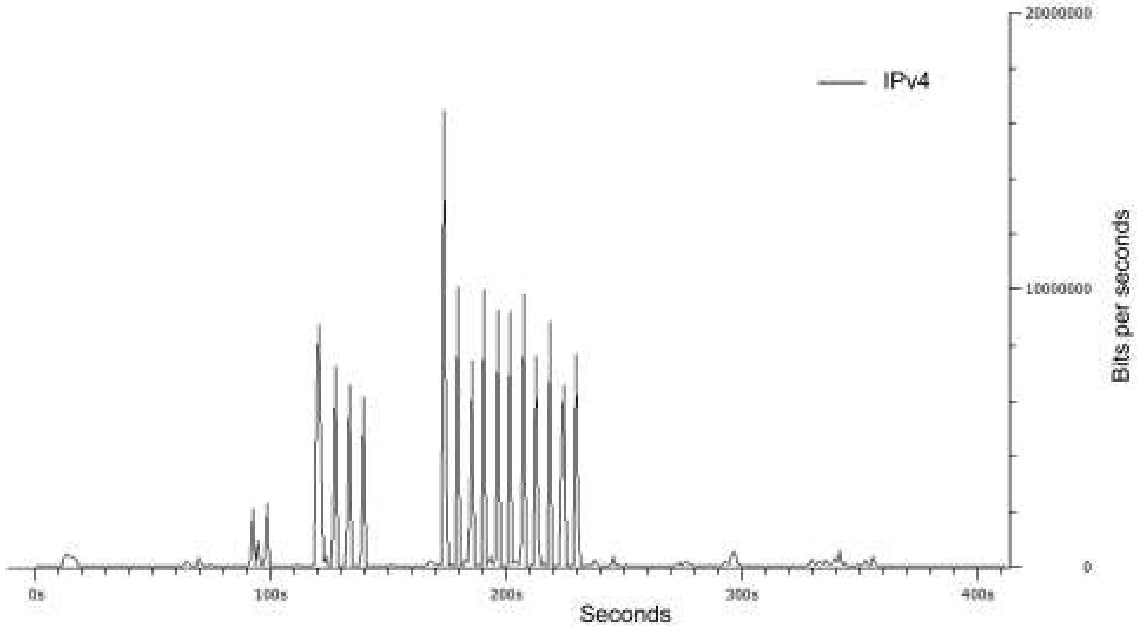

Figure 5. IPv4 bandwidth usage

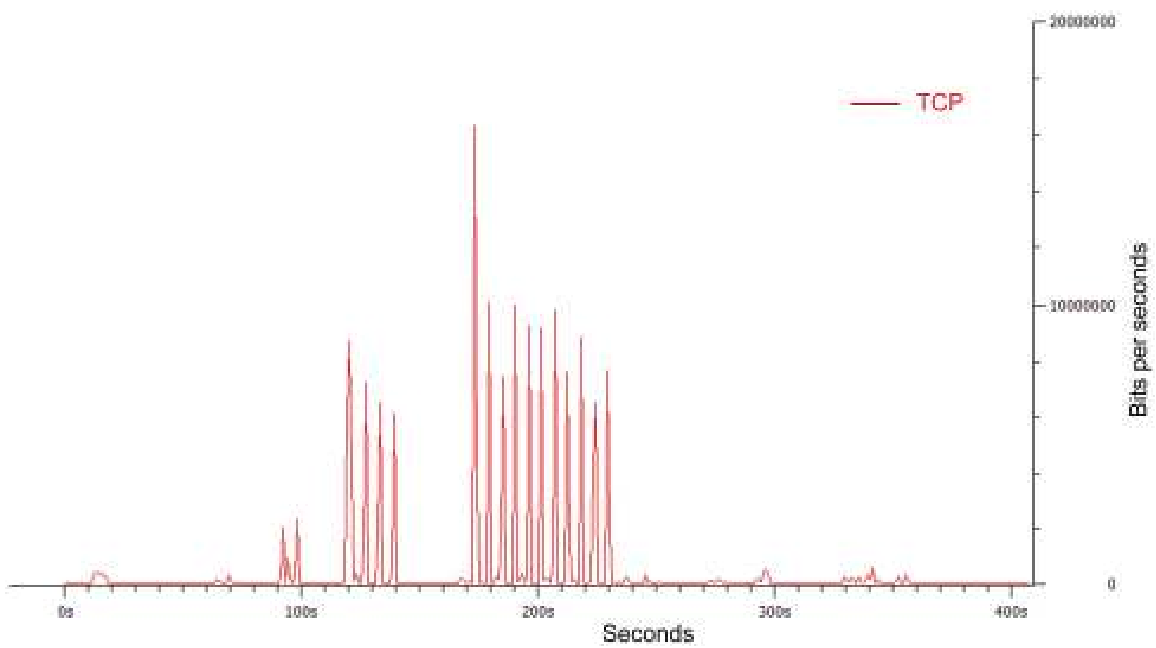

Figure 6. TCP bandwidth usage

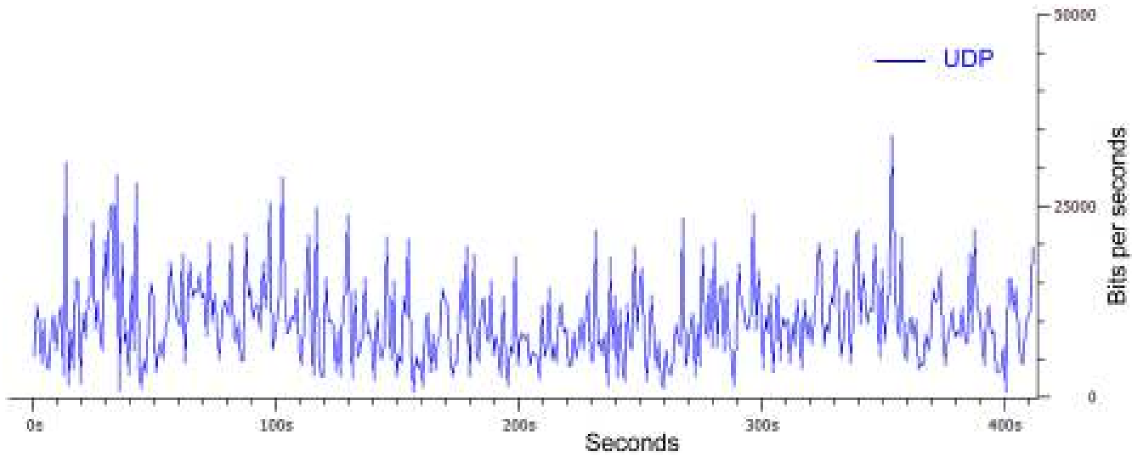

Figure 7. UDP bandwidth usage

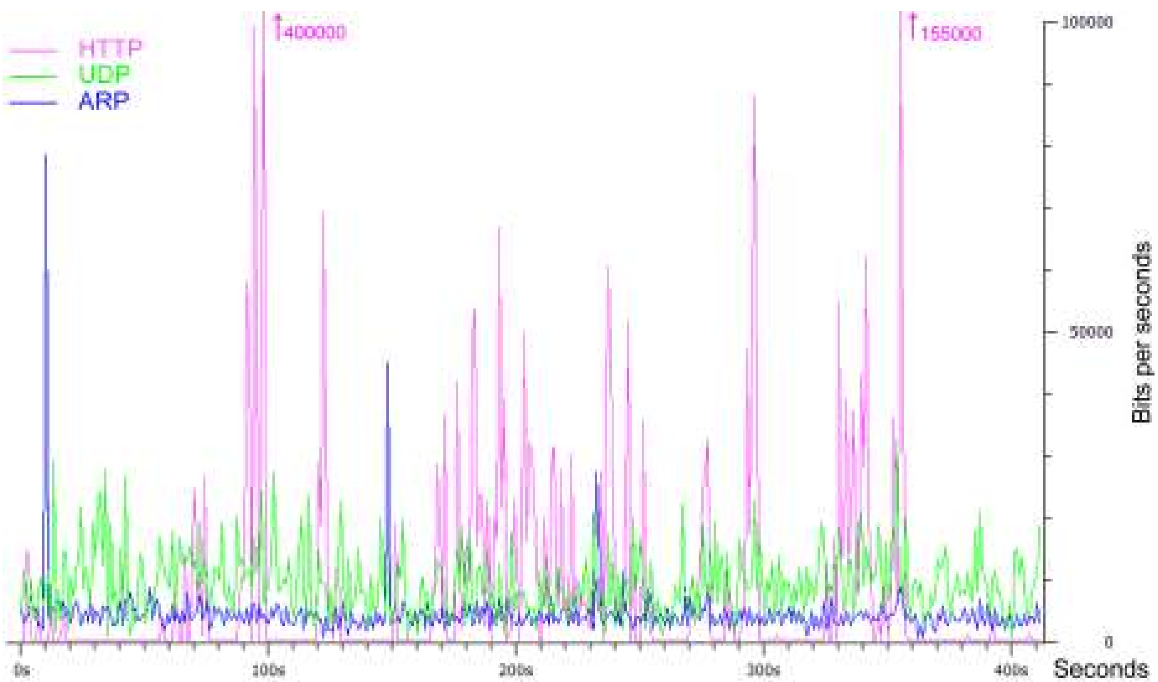

Figure 8. IPv4/TCP/UDP bandwidth usage

3.4. Confidence Interval (CI) calculation

The line traffic traces were gathered over a time period of one month. The results were analyzed by the software and the statistics were carried out. To represent the graphs we have used a time-slot, which had the closest statistical variable to the overall statistics. The 400 seconds of time period for graphs 5 and 6 can be fitted by a Gaussian model. In fact we use Gaussian approximation to calculate the confidence interval. For this we use several statistics, including those of Table 1. These statistics are gathered from different dates and daily time periods, which are s shown in Tables 2 and 3. Table 3 shows the the average number of packet sizes. Figure 9 contains packet size distribution statistics measured using the entire 10 GB of data traces. Figure 9 can be approximated well with a Gaussian distribution.

Table 2. Line traffic parameters during various time/date

Traffic Parameter	Parameter Values	Traffic Parameter	Parameter Values
Packets	32,147	Packets	19,524
Bytes	21,745,794	Bytes	15,501,610
Sniff Duration	411.829 sec	Sniff Duration	270,066 sec
Average packets/sec	78.059	Average packets/sec	72.294
Average Mbit/Sec	0.422	Average Mbit/Sec	0.459
Average bytes/sec	52,802.949	Average bytes/sec	57,399.443
Average packet size	676.449 bytes	Average packet size	793.977 bytes
Traffic Parameter	**Parameter Values**	**Traffic Parameter**	**Parameter Values**
Packets	16,956	Packets	13,794
Bytes	13,466,250	Bytes	11,507,313
Sniff Duration	429.204 sec	Sniff Duration	169.519 sec
Average packets/sec	39.506	Average packets/sec	81.371
Average Mbit/Sec	0.251	Average Mbit/Sec	0.543
Average bytes/sec	31374.942	Average bytes/sec	67,882.096
Average packet size	794.188 bytes	Average packet size	834.226 bytes
Traffic Parameter	**Parameter Values**	**Traffic Parameter**	**Parameter Values**
Packets	16,653	Packets	21,746
Bytes	13,179,228	Bytes	9,887,008
Sniff Duration	242.036 sec	Sniff Duration	499.780 sec
Average packets/sec	68.804	Average packets/sec	43.511
Average Mbit/Sec	0.436	Average Mbit/Sec	0.158
Average bytes/sec	54,451.554	Average bytes/sec	19,782.727
Average packet size	791.403 bytes	Average packet size	454.659 bytes
Traffic Parameter	**Parameter Values**	**Traffic Parameter**	**Parameter Values**
Packets	16,380	Packets	25,210
Bytes	6,052,500	Bytes	14,647,180
Sniff Duration	430.815 sec	Sniff Duration	488.747 sec
Average packets/sec	38.021	Average packets/sec	51.581
Average Mbit/Sec	0.112	Average Mbit/Sec	0.240
Average bytes/sec	14,048.944	Average bytes/sec	29,968.867
Average packet size	369.505 bytes	Average packet size	581.007 bytes

Table 3. Average packet size statistics

Packet Size (Bytes)	Number of Packets
676.449	32,147
793.977	19,524
794.188	16,956
834.226	13,794
791.403	16,653
454.659	21,746
369.505	16,380
581.007	25,210

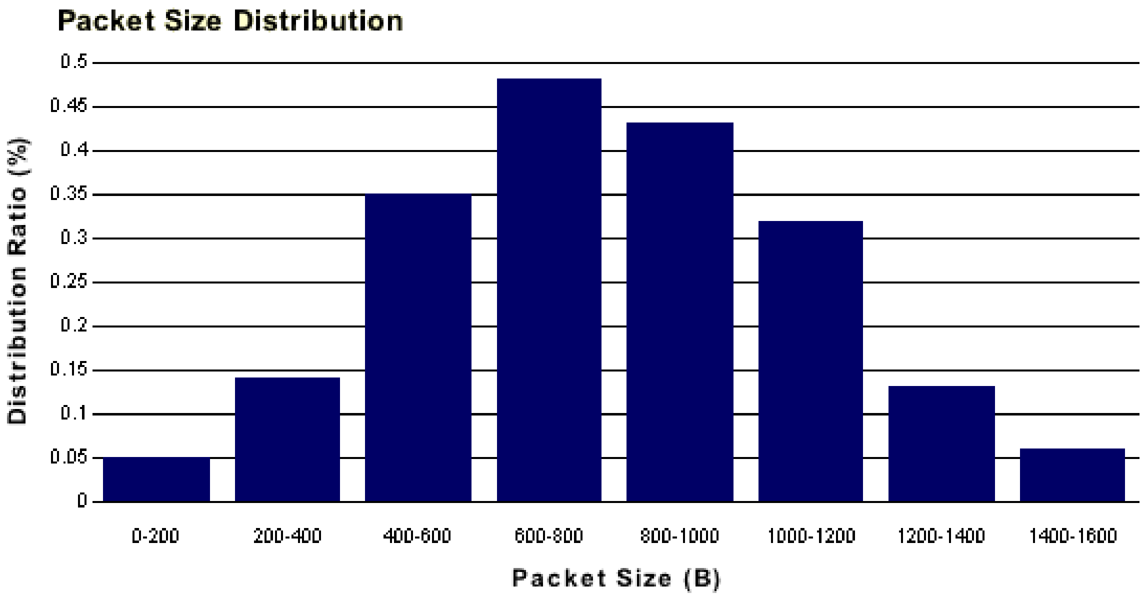

Figure 9. Packet size distributions

The average packet sizes, variance, and standard deviation values can be calculated from Table 3 data. There are a total number of 162,410 packets containing over 105,986,883 bytes. The total average packet size = 652.588 B (bytes), and Standard Deviation (STD) = 153 bytes, and the average bytes per packet is 652 ± 1 bytes. The same calculation is carried out for the flow durations, which results 384 sec average flow durations.

Therefore this leads to the following value for the CI:

$$CI \approx 2.57 \times 153.33 \div 403 = 0.9778$$

Therefore for the byte size interval of 652.588 ± 0.9778 bytes we can achieve a 99% confidence.

3.5. TCP traffic versus UDP traffic

The data analysis of the wired traffic presented in section III showed that the majority of the traffic was based on TCP-based (e.g., HTTP) connections. In this section, we analyze a wired link where the majority of the traffic is based on UDP streaming data. For this we have captured more than 5 GB of data. The following statistics are observed: IPv4 (%96.91), UDP (%91.50), and TCP (%4.39). The average package size is 594.45 ± 1 byes (99% CI) with an average connection time of 365.24 seconds (Table 4). The results are very close to the results presented in section III. The difference is mostly due to the difference in the TCP/UDP header sizes.

Table 4. Average statistics for wireless and wired flows

Average Number of bytes/packet	Average Flow Duration (Sec)
652.588 (TCP Dominant Traffic)	384.24 (TCP Dominant Traffic)
594.45 (UDP Dominant Traffic)	365.24 (UDP Dominant Traffic)

4. Conclusions

Data mining has been discussed in this paper, based on a few traffic classification approaches, including: Packet size and flow duration. Then heavy hitters (long lived flows and elephant/dragonflies packets) and light hitters (tortoises flows or mice packets) were introduced. Basically 20% of the connections on the Internet are responsible for 80% of the traffic passing through a single node, mostly including elephant packets.

Bandwidth fluctuations (in particular in wireless networks) were found to be mostly due sudden link exhaustion, connection termination, burstiness, and RF issues. However in wired networks, fluctuations were mostly due to sudden changes in the client bandwidth demands and congestion periods.

Then we considered monitoring a wired link and sniffing wired-based traffic for a period of one month using Wireshark software. The monitoring sessions took place during four time periods: 8 AM – 9 AM, 12 PM – 1 PM, 4 PM – 5 PM, and 12 AM – 1 AM. Scatter plot graphs were used to represents the related statistical values for the duration time and transmitted bytes.

We believe the selection of these four time periods proves to provide a diverse and random selection of packets, suitable for our calculations. We calculated the average packet size and flow duration. For this we used confidence intervals, which provide estimates on how reliable our samples are.

Throughput graphs for IPv4, TCP, and UDP were also depicted. Table XXX showed that the average number of bytes per packet for the TCP-based traffic was approximately 652 bytes and for UDP-based traffic was found to be 594 bytes. The average flow duration for TCP-based traffic was calculated to be 384 seconds and for UDP-based traffic flows was 365 seconds.

5. References

[1] Jörg Wallerich, Holger Dreger, Anja Feldmann, Balachander Krishnamurthy, Walter Willinger. A methodology for studying persistency aspects of internet flows. SIGCOMM Computer Communications Review (CCR), 35(2):23-36, 2005

[2] Jörg Wallerich, "Capturing the Variability of Internet Flows in a Workload Generator for Network Simulators", Ph.D. Dissertation, University of Munich, 2007

[3] Andre Broido, Young Hyun, Ruomei Gao, and kc claffy, "Their share: diversity and disparity in ip traffic". In The 5th anuual Passive & Active Measurement Workshop, PAM2004 (LNCS3015), 2004.

[4] Wentian Li. References on zipf's law, http://linkage.rockefeller.edu/wli/zipf/, 2003

[5] Konstantina Papagiannaki, Nina Taft, Christophe Diot, "Impact of Flow Dynamics on Traffic Engineering Design Principles", INFOCOM 2004, 7-11 March 2004, Vol. 4, Page(s): 2295- 2306

[6] J. Cao, W.S. Cleveland, Y. Gao, K. Jeffay, F.D. Smith, and M.C. Weigle, Stochastic Models for Generating Synthetic HTTP Source Traffic, IEEE INFOCOM, March 2004

[7] IX SYMPOSIUM ON PROBABILITY AND STOCHASTIC PROCESSES, November 20-24, 2006 CIMAT, Guanajuato Gto

[8] J. P. Dubois, Burstiness Reduction of a Doubly Stochastic AR-Modeled Uniform Activity VBR Video, Proc. WASET, 23: 454-458, 2007

[9] Kun-Chan Lan, John Heidemann, "A measurement study of correlations of Internet flow characteristics", USC Information Sciences Institute, ISI-TR-574

[10] Yin Zhang, Lee Breslau, Vern Paxson, and Scott Shenker, "On the characteristics and origins of internet flow rates. In SIGCOMM, Pittsburgh, PA, USA, August 2002

[11] Andrew W. Moore and Konstantina Papagiannaki, "Toward the Accurate Identification of Network Applications", Passive and Active Measurements Workshop, Boston, MA, USA, March 31 - April 1, 2005.

[12] Thomas Karagiannis, Andre Broido, Nevil Brownlee, kc claffy, “Is P2P dying or just hiding?”, Proceedings of Globecom 2004, November/December 2004.

[13] Alok Madhukar, Carey L. Williamson, “A Longitudinal Study of P2P Traffic Classification”, MASCOTS 2006

[14] Wireshark, http://www.wireshark.org/

[15] Scatter Plot Tool, Mathcracker, http://www.mathcracker.com/scatter_plot.php

Appendix A

Table A.1 Protocol Hierarchy

Protocol	% Packets	Packets	Bytes	Mbit/s	End Packets	End Bytes	End Mbit/s
⊟ Frame	100.00%	32147	21745794	0.422	0	0	0.000
⊟ Ethernet	99.96%	32133	21744534	0.422	0	0	0.000
⊟ Internet Protocol	87.62%	28166	21492336	0.418	0	0	0.000
⊟ User Datagram Protocol	8.36%	2686	478999	0.009	47	3290	0.000
Domain Name Service	1.00%	323	46514	0.001	323	46514	0.001
⊟ NetBIOS Datagram Service	2.59%	833	202419	0.004	0	0	0.000
⊟ SMB (Server Message Block Protocol)	2.59%	833	202419	0.004	0	0	0.000
⊟ SMB MailSlot Protocol	2.59%	833	202419	0.004	0	0	0.000
Microsoft Windows Browser Protocol	2.56%	824	199732	0.004	824	199732	0.004
Microsoft Windows Logon Protocol (Old)	0.03%	9	2687	0.000	9	2687	0.000
Data	0.70%	225	34065	0.001	225	34065	0.001
NetBIOS Name Service	1.61%	518	48589	0.001	518	48589	0.001
Cisco Hot Standby Router Protocol	1.35%	435	26942	0.001	435	26942	0.001
Session Initiation Protocol	0.00%	1	582	0.000	1	582	0.000
Bootstrap Protocol	0.75%	242	105521	0.002	242	105521	0.002
Routing Information Protocol	0.04%	14	1204	0.000	14	1204	0.000
Connectionless Lightweight Directory Access Protocol	0.13%	42	9387	0.000	42	9387	0.000
⊟ UDP Encapsulation of IPsec Packets	0.01%	4	306	0.000	3	180	0.000
Encapsulating Security Payload	0.00%	1	126	0.000	1	126	0.000
Network Time Protocol	0.01%	2	180	0.000	2	180	0.000
⊟ Transmission Control Protocol	78.81%	25335	20997909	0.408	24503	20366689	0.396
⊟ Hypertext Transfer Protocol	1.59%	511	296653	0.006	374	219475	0.004
Line-based text data	0.21%	66	35873	0.001	66	35873	0.001
Compuserve GIF	0.16%	53	24416	0.000	53	24416	0.000
JPEG File Interchange Format	0.01%	3	3650	0.000	3	3650	0.000
Media Type	0.02%	5	4340	0.000	5	4340	0.000
eXtensible Markup Language	0.02%	7	5053	0.000	7	5053	0.000
Portable Network Graphics	0.01%	3	3846	0.000	3	3846	0.000
TPKT - ISO on TCP - RFC1006	0.02%	5	2060	0.000	5	2060	0.000
Secure Socket Layer	0.97%	313	331534	0.006	313	331534	0.006
Data	0.01%	3	973	0.000	3	973	0.000
Internet Control Message Protocol	0.30%	97	11024	0.000	97	11024	0.000
Open Shortest Path First	0.13%	41	3690	0.000	41	3690	0.000
Generic Routing Encapsulation	0.02%	7	714	0.000	7	714	0.000
Address Resolution Protocol	11.38%	3658	222710	0.004	3658	222710	0.004
⊟ Logical-Link Control	0.76%	245	17132	0.000	16	960	0.000
Spanning Tree Protocol	0.64%	205	12300	0.000	205	12300	0.000
Dynamic Trunking Protocol	0.04%	14	840	0.000	14	840	0.000
Cisco Discovery Protocol	0.02%	7	2828	0.000	7	2828	0.000
Logical-Link Control Basic Format XID	0.00%	1	60	0.000	1	60	0.000
Cisco Wireless LAN Context Control Protocol	0.01%	2	144	0.000	2	144	0.000
⊟ Configuration Test Protocol (loopback)	0.13%	43	2580	0.000	0	0	0.000
Data	0.13%	43	2580	0.000	43	2580	0.000
ATAoverEthernet	0.02%	7	8568	0.000	7	8568	0.000
Cisco Wireless LAN Context Control Protocol	0.00%	1	60	0.000	1	60	0.000
⊟ Internet Protocol Version 6	0.02%	7	686	0.000	0	0	0.000
Internet Control Message Protocol v6	0.01%	4	344	0.000	4	344	0.000
⊟ User Datagram Protocol	0.01%	3	342	0.000	0	0	0.000
DHCPv6	0.01%	3	342	0.000	3	342	0.000
Data	0.02%	6	462	0.000	6	462	0.000
⊟ Cisco ISL	0.04%	14	1260	0.000	0	0	0.000
⊟ Ethernet	0.04%	14	1260	0.000	0	0	0.000
⊟ Logical-Link Control	0.04%	14	1260	0.000	0	0	0.000
Dynamic Trunking Protocol	0.04%	14	1260	0.000	14	1260	0.000

Table A.2 Ethernet flows ordered by heavy hitter end point bytes

Ethernet Conversations

Address A	Address B	Packets	Bytes	Packets A->B	Bytes A->B	Packets A<-B	Bytes A<-B	Rel Start	Duration	bps A->B	bps A<-B
xxxx	yyyy	15340	20079465	15340	20079465	0	0	2.381945000	404.9778	396653.12	N/A
xxxx	yyyy	10272	948450	0	0	10272	948450	1.801738000	405.7396	N/A	18700.66
xxxx	yyyy	2262	135935	2262	135935	0	0	0.116182000	411.4847	2642.82	N/A
xxxx	yyyy	80	47200	80	47200	0	0	32.577021000	376.9135	1001.82	N/A
xxxx	yyyy	172	29274	82	6541	90	22733	0.000000000	358.1720	146.10	507.76
xxxx	yyyy	209	21359	209	21359	0	0	53.326330000	358.1894	477.04	N/A
xxxx	yyyy	81	19598	81	19598	0	0	13.074541000	398.3006	393.63	N/A
xxxx	yyyy	86	18170	86	18170	0	0	13.074579000	398.2852	364.96	N/A
xxxx	yyyy	257	17350	257	17350	0	0	0.011066000	409.9104	338.61	N/A
xxxx	yyyy	224	13860	224	13860	0	0	1.592621000	408.4631	271.46	N/A
xxxx	yyyy	211	13082	211	13082	0	0	1.867859000	409.6521	255.48	N/A
xxxx	yyyy	205	12300	205	12300	0	0	0.103039000	411.0511	239.39	N/A
xxxx	yyyy	62	10332	62	10332	0	0	10.819365000	394.9700	209.27	N/A
xxxx	yyyy	58	10306	58	10306	0	0	7.070626000	401.4037	205.40	N/A
xxxx	yyyy	8	8660	8	8660	0	0	17.096301000	359.9898	192.45	N/A
xxxx	yyyy	114	7475	114	7475	0	0	3.449150000	309.2908	193.35	N/A
xxxx	yyyy	47	7427	47	7427	0	0	15.766646000	384.0010	154.73	N/A
xxxx	yyyy	43	7155	43	7155	0	0	2.139889000	399.6688	143.22	N/A
xxxx	yyyy	33	6587	33	6587	0	0	10.439151000	399.8668	131.78	N/A
xxxx	yyyy	35	6582	35	6582	0	0	43.387229000	342.7116	153.65	N/A
xxxx	yyyy	29	6530	29	6530	0	0	10.836523000	400.2669	130.51	N/A
xxxx	yyyy	29	6530	29	6530	0	0	4.457948000	402.3805	129.83	N/A
xxxx	yyyy	28	6470	28	6470	0	0	11.701854000	400.0232	129.39	N/A
xxxx	yyyy	28	6470	28	6470	0	0	1.811931000	399.9830	129.41	N/A
xxxx	yyyy	27	6410	27	6410	0	0	7.543851000	400.0421	128.19	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.420542000	400.0389	128.19	N/A
xxxx	yyyy	27	6410	27	6410	0	0	4.265939000	400.2831	128.11	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.083161000	400.2283	128.13	N/A
xxxx	yyyy	27	6410	27	6410	0	0	9.768650000	400.2297	128.13	N/A
xxxx	yyyy	27	6410	27	6410	0	0	11.563555000	400.2656	128.11	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.945987000	400.2367	128.12	N/A
xxxx	yyyy	26	6318	26	6318	0	0	4.477096000	400.3840	126.24	N/A
xxxx	yyyy	26	6318	26	6318	0	0	5.872089000	400.0094	126.36	N/A
xxxx	yyyy	26	6318	26	6318	0	0	8.202490000	400.2736	126.27	N/A
xxxx	yyyy	26	6318	26	6318	0	0	9.814400000	400.2656	126.28	N/A
xxxx	yyyy	26	6318	26	6318	0	0	10.294794000	400.0389	126.35	N/A
xxxx	yyyy	26	6318	26	6318	0	0	1.542215000	400.0226	126.35	N/A
xxxx	yyyy	27	6227	27	6227	0	0	4.809094000	399.6452	124.65	N/A
xxxx	yyyy	26	6167	26	6167	0	0	13.719238000	384.3963	128.35	N/A
xxxx	yyyy	25	6075	25	6075	0	0	0.000282000	399.6341	121.61	N/A
xxxx	yyyy	25	6075	25	6075	0	0	1.986426000	399.7799	121.57	N/A
xxxx	yyyy	87	6014	85	5930	2	84	2.106739000	405.2319	117.07	1.66
xxxx	yyyy	25	5924	25	5924	0	0	13.834749000	383.4658	123.59	N/A
xxxx	yyyy	36	5922	36	5922	0	0	17.396969000	348.0261	136.13	N/A
xxxx	yyyy	36	5278	36	5278	0	0	93.307040000	277.9280	151.92	N/A
xxxx	yyyy	35	4928	35	4928	0	0	12.172357000	399.0056	98.81	N/A
xxxx	yyyy	28	4500	28	4500	0	0	10.558705000	343.3400	104.85	N/A
xxxx	yyyy	47	4155	47	4155	0	0	13.504966000	378.1107	87.91	N/A
xxxx	yyyy	23	3750	23	3750	0	0	0.120452000	364.7750	82.24	N/A
xxxx	yyyy	41	3690	41	3690	0	0	5.138473000	400.0161	73.80	N/A
xxxx	yyyy	19	3396	19	3396	0	0	29.060473000	364.3001	74.58	N/A
xxxx	yyyy	37	3354	37	3354	0	0	24.625805000	320.9413	83.60	N/A
xxxx	yyyy	16	3002	16	3002	0	0	12.221941000	390.9727	61.43	N/A
xxxx	yyyy	42	2867	42	2867	0	0	14.386653000	391.3686	58.60	N/A
xxxx	yyyy	28	2865	28	2865	0	0	43.521578000	364.0251	62.96	N/A
xxxx	yyyy	2	2828	2	2828	0	0	21.251742000	0.0001	209481481.48	N/A
xxxx	yyyy	41	2803	41	2803	0	0	15.891203000	391.2518	57.31	N/A
xxxx	yyyy	42	2688	42	2688	0	0	2.675875000	400.1701	53.74	N/A
xxxx	yyyy	13	2472	13	2472	0	0	154.424451000	121.8761	162.26	N/A
xxxx	yyyy	41	2460	41	2460	0	0	7.895929000	400.2255	49.17	N/A
xxxx	yyyy	26	2431	26	2431	0	0	0.999890000	386.2026	50.36	N/A
xxxx	yyyy	21	2210	21	2210	0	0	83.998103000	312.7562	56.53	N/A
xxxx	yyyy	15	1967	15	1967	0	0	55.778928000	299.7265	52.50	N/A

Table A.3 IPv4 flows ordered by heavy hitter end point bytes

IPv4 Conversations

Address A	Address B	Packets	Bytes	Packets A->B	Bytes A->B	Packets A<-B	Bytes A<-B	Rel Start	Duration	bps A->B	bps A<-B
xxxx	yyyy	14978	13576738	5690	312084	9288	13264654	151.857165000	78.5998	31764.35	1350095.03
xxxx	yyyy	5452	4981695	2033	114741	3419	4866954	97.821744000	52.8085	17382.19	737298.26
xxxx	yyyy	384	244225	142	14700	242	229525	3.786964000	403.7544	291.27	4547.81
xxxx	yyyy	364	74148	217	34285	147	39863	72.234699000	173.3670	1582.08	1839.47
xxxx	yyyy	352	297661	133	11161	219	286500	70.682240000	86.1442	1036.49	26606.55
xxxx	yyyy	343	266270	137	21051	206	245219	97.714167000	157.6042	1068.55	12447.34
xxxx	yyyy	323	89897	169	27170	154	62727	176.160782000	78.8098	2758.03	6367.43
xxxx	yyyy	231	66901	126	51844	105	15057	71.684330000	225.1206	1842.35	535.07
xxxx	yyyy	224	13860	224	13860	0	0	1.592621000	408.4631	271.46	N/A
xxxx	yyyy	212	92378	105	52006	107	40372	277.289360000	64.5930	6441.07	5000.17
xxxx	yyyy	211	13082	211	13082	0	0	1.867859000	409.6521	255.48	N/A
xxxx	yyyy	209	21359	209	21359	0	0	53.326330000	358.1694	477.04	N/A
xxxx	yyyy	189	163250	73	5906	116	157344	123.655123000	131.6629	358.86	9560.41
xxxx	yyyy	172	29274	90	22733	82	6541	0.000000000	358.1720	507.76	146.10
xxxx	yyyy	157	130811	61	11257	96	119554	167.698619000	87.6200	1027.80	10915.68
xxxx	yyyy	152	17247	74	11243	78	6004	1.801738000	351.9408	255.57	136.48
xxxx	yyyy	140	68234	140	68234	0	0	13.073987000	398.3008	1370.50	N/A
xxxx	yyyy	123	51617	62	15434	61	36183	213.084099000	194.2759	635.55	1489.96
xxxx	yyyy	121	34499	66	23313	55	11186	170.964937000	72.7614	2563.23	1229.88
xxxx	yyyy	120	38938	59	13785	61	25153	90.437917000	7.3834	14936.26	27253.67
xxxx	yyyy	108	26385	57	12259	51	14126	271.934199000	24.3769	4023.16	4635.87
xxxx	yyyy	107	54917	52	14817	55	40100	275.976091000	22.0299	5380.68	14562.01
xxxx	yyyy	104	78132	41	5186	63	72946	97.888721000	63.9374	648.89	9127.18
xxxx	yyyy	97	71158	39	7915	58	63243	332.161318000	7.0007	9044.83	72270.65
xxxx	yyyy	92	72425	37	5045	55	67380	94.368076000	0.3796	106335.61	1420196.92
xxxx	yyyy	86	61593	38	6119	48	55474	64.721363000	271.1217	180.55	1636.87
xxxx	yyyy	86	49995	38	10697	48	39298	353.743737000	12.5070	6842.26	25136.69
xxxx	yyyy	83	5810	83	5810	0	0	2.106739000	405.2319	114.70	N/A
xxxx	yyyy	80	25251	44	12418	36	12833	169.088048000	76.3507	1301.15	1344.64
xxxx	yyyy	79	22181	45	9221	34	12960	273.425116000	57.7239	1277.95	1796.14
xxxx	yyyy	76	53423	30	4280	46	49143	340.589088000	65.6618	521.46	5987.41
xxxx	yyyy	75	43817	35	7068	40	36749	69.202014000	97.4208	580.41	3017.75
xxxx	yyyy	75	14416	43	8738	32	5678	292.124201000	60.7217	1151.22	748.07
xxxx	yyyy	64	14913	35	11007	29	3906	2.317305000	94.5296	931.52	330.56
xxxx	yyyy	58	13147	32	9135	26	4012	66.090441000	265.8444	274.90	120.73
xxxx	yyyy	54	36031	23	2485	31	33546	296.458490000	99.7920	199.21	2689.27
xxxx	yyyy	54	3204	0	0	54	3204	139.741283000	228.5816	N/A	112.14
xxxx	yyyy	49	15534	24	5854	25	9680	275.410645000	20.1225	2327.35	3848.43
xxxx	yyyy	47	17343	47	17343	0	0	13.074541000	398.3006	348.34	N/A
xxxx	yyyy	47	3290	47	3290	0	0	280.502317000	24.5830	1070.66	N/A
xxxx	yyyy	45	29156	19	2932	26	26224	349.526051000	16.7250	1402.45	12543.65
xxxx	yyyy	44	13686	24	10818	20	2868	334.160841000	7.1016	12186.55	3230.82
xxxx	yyyy	43	36042	16	1946	27	34096	297.743438000	64.1394	242.72	4252.73
xxxx	yyyy	42	15498	42	15498	0	0	17.387569000	374.9915	330.63	N/A
xxxx	yyyy	41	3690	41	3690	0	0	5.138473000	400.0161	73.80	N/A
xxxx	yyyy	38	9439	22	3507	16	5932	62.416601000	303.8346	92.34	156.19
xxxx	yyyy	37	3555	37	3555	0	0	13.504966000	378.1107	75.22	N/A
xxxx	yyyy	36	5922	36	5922	0	0	17.396969000	348.0261	136.13	N/A
xxxx	yyyy	36	3312	36	3312	0	0	24.625805000	320.9413	82.56	N/A
xxxx	yyyy	33	12399	17	6930	16	5469	279.008618000	16.7104	3317.70	2618.25
xxxx	yyyy	32	5314	32	5314	0	0	10.819365000	394.9700	107.63	N/A
xxxx	yyyy	32	5022	32	5022	0	0	101.252798000	269.9822	148.81	N/A
xxxx	yyyy	32	8119	16	3262	16	4857	122.308613000	0.2266	115170.40	171484.56
xxxx	yyyy	31	6964	11	3147	20	3817	69.294935000	45.2347	556.56	675.06
xxxx	yyyy	31	3001	16	1723	15	1278	270.664840000	30.4613	452.51	335.64
xxxx	yyyy	30	5018	30	5018	0	0	32.914875000	370.4624	108.36	N/A
xxxx	yyyy	30	5286	30	5286	0	0	7.070626000	401.4037	105.35	N/A
xxxx	yyyy	30	6282	30	6282	0	0	55.721919000	330.3769	152.12	N/A
xxxx	yyyy	30	7390	15	4301	15	3089	96.670488000	55.2008	623.32	447.67
xxxx	yyyy	28	5020	28	5020	0	0	61.595281000	342.7364	117.17	N/A
xxxx	yyyy	27	6410	27	6410	0	0	7.543851000	400.0421	128.19	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.420542000	400.0389	128.19	N/A
xxxx	yyyy	27	6410	27	6410	0	0	4.265939000	400.2831	128.11	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.083161000	400.2283	128.13	N/A
xxxx	yyyy	27	6410	27	6410	0	0	9.768650000	400.2297	128.13	N/A
xxxx	yyyy	27	6410	27	6410	0	0	11.563555000	400.2656	128.11	N/A
xxxx	yyyy	27	6410	27	6410	0	0	8.945987000	400.2367	128.12	N/A
xxxx	yyyy	27	6410	27	6410	0	0	4.457948000	402.3805	127.44	N/A
xxxx	yyyy	27	6410	27	6410	0	0	10.836523000	400.2669	128.11	N/A

Table A.4 TCP flows ordered by heavy hitter end point bytes

TCP Conversations

Address A	Port A	Address B	Port B	Packets	Bytes	Packets A->B	Bytes A->B	Packets A<-B	Bytes A<-B	Rel Start	Duration	bps A->B	bps A<-B
xxxx	yyyy	xxxx	yyyy	14975	13576552	5687	311898	9288	13264654	172.935077000	57.5219	43377.97	1844814.06
xxxx	yyyy	xxxx	yyyy	5449	4981509	2030	114555	3419	4866954	118.699987000	31.9303	28701.28	1219395.26
xxxx	yyyy	xxxx	yyyy	374	241991	136	13005	238	228986	4.976367000	402.5650	258.44	4550.54
xxxx	yyyy	xxxx	yyyy	190	168550	70	5083	120	163467	91.870863000	64.9556	626.03	20132.78
xxxx	yyyy	xxxx	yyyy	147	128965	56	4550	91	124415	193.914975000	61.4033	592.80	16209.61
xxxx	yyyy	xxxx	yyyy	144	124230	52	4537	92	119693	70.682240000	86.1441	421.34	11115.61
xxxx	yyyy	xxxx	yyyy	141	110039	56	9150	85	100889	97.714167000	64.1120	1141.75	12589.10
xxxx	yyyy	xxxx	yyyy	144	108574	59	10046	85	98528	97.714180000	64.1120	1253.56	12294.48
xxxx	yyyy	xxxx	yyyy	98	81126	38	7365	60	73761	167.698619000	62.3375	945.18	9466.03
xxxx	yyyy	xxxx	yyyy	63	56627	23	2257	40	54370	94.534794000	0.2128	84835.67	2043648.84
xxxx	yyyy	xxxx	yyyy	59	49685	23	3892	36	45793	236.924035000	18.3946	1692.67	19915.88
xxxx	yyyy	xxxx	yyyy	58	47657	22	1855	36	45802	184.930031000	70.3883	210.83	5205.64
xxxx	yyyy	xxxx	yyyy	52	45001	20	2177	32	42824	341.899464000	64.3514	270.64	5323.77
xxxx	yyyy	xxxx	yyyy	59	44201	23	2936	36	41265	97.888721000	63.9374	367.36	5163.18
xxxx	yyyy	xxxx	yyyy	45	37443	18	2043	27	35400	69.202014000	0.2218	73679.70	1276682.07
xxxx	yyyy	xxxx	yyyy	43	36042	16	1946	27	34096	297.740438000	64.1394	242.72	4252.73
xxxx	yyyy	xxxx	yyyy	42	34285	17	1356	25	32929	123.695123000	62.6054	173.28	4207.91
xxxx	yyyy	xxxx	yyyy	45	33931	18	2250	27	31681	97.888750000	63.9373	281.53	3964.01
xxxx	yyyy	xxxx	yyyy	52	32184	22	6214	30	25970	353.740737000	12.5068	3974.78	16611.71
xxxx	yyyy	xxxx	yyyy	39	31091	15	3083	24	28008	338.578703000	0.5833	42283.63	384132.32
xxxx	yyyy	xxxx	yyyy	39	29911	17	2512	22	27399	329.247670000	0.3817	52646.33	574226.46
xxxx	yyyy	xxxx	yyyy	38	29494	16	2178	22	27316	64.721363000	0.2932	59434.92	745419.75
xxxx	yyyy	xxxx	yyyy	45	29156	19	2932	26	26224	349.526051000	16.7250	1402.45	12543.65
xxxx	yyyy	xxxx	yyyy	128	28901	64	5168	64	23733	214.697490000	22.2290	1859.91	8541.26
xxxx	yyyy	xxxx	yyyy	35	27934	14	2827	21	25107	295.017687000	0.3574	63270.56	561915.13
xxxx	yyyy	xxxx	yyyy	32	22558	13	2594	19	19964	335.587790000	0.4276	48534.74	373534.16
xxxx	yyyy	xxxx	yyyy	35	20210	16	3866	19	16344	213.084099000	39.4189	784.60	3316.99
xxxx	yyyy	xxxx	yyyy	27	17997	12	2226	15	15771	342.389444000	15.0018	1187.06	8410.21
xxxx	yyyy	xxxx	yyyy	34	17811	16	4483	18	13328	355.594118000	10.6566	3365.43	10005.45
xxxx	yyyy	xxxx	yyyy	26	17509	11	2238	15	15271	332.161318000	0.6379	28067.40	191518.01
xxxx	yyyy	xxxx	yyyy	22	16106	9	892	13	15214	296.459181000	0.1765	40437.47	689703.63
xxxx	yyyy	xxxx	yyyy	40	14214	20	2828	20	11386	176.160782000	6.0781	3722.21	14986.23
xxxx	yyyy	xxxx	yyyy	20	14047	8	838	12	13209	296.459490000	0.0249	268966.90	4239598.80
xxxx	yyyy	xxxx	yyyy	20	13770	9	1502	11	12268	94.422307000	0.1344	89437.37	730504.42
xxxx	yyyy	xxxx	yyyy	19	10935	10	6792	9	4143	296.462272000	0.7906	68726.77	41922.11
xxxx	yyyy	xxxx	yyyy	20	9555	9	3164	11	6391	296.480988000	0.4594	55102.15	111301.47
xxxx	yyyy	xxxx	yyyy	34	9470	18	7010	16	2460	2.317876000	42.1383	1330.86	467.03
xxxx	yyyy	xxxx	yyyy	16	9365	8	2026	8	7339	273.791487000	0.3172	51091.46	185074.16
xxxx	yyyy	xxxx	yyyy	17	9221	8	2023	9	7198	275.410649000	0.6333	25556.72	90932.92
xxxx	yyyy	xxxx	yyyy	17	8968	8	1670	9	7298	95.604717000	0.1929	69271.97	302722.66
xxxx	yyyy	xxxx	yyyy	17	8830	8	3937	9	4893	295.051322000	0.3643	86447.44	107438.99
xxxx	yyyy	xxxx	yyyy	15	8575	7	1055	8	7520	203.980446000	0.2527	33393.47	238027.41
xxxx	yyyy	xxxx	yyyy	15	8571	7	1055	8	7516	183.376380000	0.3000	28134.08	200432.01
xxxx	yyyy	xxxx	yyyy	15	8568	7	1058	8	7510	193.411856000	0.1414	59839.51	424758.74
xxxx	yyyy	xxxx	yyyy	30	8092	16	2564	14	5528	251.382074000	3.5885	5715.97	12323.66
xxxx	yyyy	xxxx	yyyy	16	7926	8	1979	8	5947	341.777088000	0.1309	120920.50	363372.51
xxxx	yyyy	xxxx	yyyy	15	7904	7	1179	8	6725	296.469743000	0.4109	22952.59	130921.27
xxxx	yyyy	xxxx	yyyy	15	7876	7	1300	8	6576	91.659894000	0.1418	73365.50	371116.57
xxxx	yyyy	xxxx	yyyy	16	7607	8	4315	8	3292	339.079871000	0.5568	62001.36	47302.08
xxxx	yyyy	xxxx	yyyy	16	7599	8	4318	8	3281	339.176849000	0.4655	74203.60	56383.05
xxxx	yyyy	xxxx	yyyy	16	7347	8	3938	8	3409	340.968094000	0.9142	34459.19	29830.21
xxxx	yyyy	xxxx	yyyy	16	7065	8	3952	8	3113	340.385548000	0.4032	78412.50	61765.72
xxxx	yyyy	xxxx	yyyy	16	7027	8	3937	8	3090	340.930136000	0.3761	83749.70	65731.92
xxxx	yyyy	xxxx	yyyy	20	6963	10	3859	10	3104	279.008618000	7.5651	4080.84	3282.44
xxxx	yyyy	xxxx	yyyy	16	6860	8	3913	8	2947	332.936489000	1.0874	28787.54	21680.78
xxxx	yyyy	xxxx	yyyy	16	6820	8	3952	8	2868	335.948481000	0.4801	65847.19	47785.86
xxxx	yyyy	xxxx	yyyy	16	6805	8	3936	8	2869	293.770161000	0.3325	94687.94	69019.23
xxxx	yyyy	xxxx	yyyy	16	6804	8	3936	8	2868	277.289375000	0.3428	91850.49	66927.64
xxxx	yyyy	xxxx	yyyy	16	6742	8	3912	8	2830	332.907731000	1.0988	28481.62	20604.03
xxxx	yyyy	xxxx	yyyy	26	6606	14	2401	12	4205	238.971263000	4.6793	4104.87	7189.08
xxxx	yyyy	xxxx	yyyy	16	6343	8	3949	8	2394	335.924178000	0.5095	62009.66	37592.08
xxxx	yyyy	xxxx	yyyy	16	6343	8	3949	8	2394	340.376785000	0.4309	73320.06	44448.78
xxxx	yyyy	xxxx	yyyy	12	5878	6	755	6	5123	333.912842000	62.3377	96.89	657.45
xxxx	yyyy	xxxx	yyyy	14	5806	7	772	7	5034	352.739301000	13.5115	457.08	2980.57
xxxx	yyyy	xxxx	yyyy	13	5772	7	2713	6	3059	123.968400000	62.2920	348.42	392.86
xxxx	yyyy	xxxx	yyyy	12	5535	6	1456	6	4079	276.539183000	0.1290	90280.58	252922.03
xxxx	yyyy	xxxx	yyyy	16	5522	9	1721	7	3801	348.505531000	17.7457	775.85	1713.54
xxxx	yyyy	xxxx	yyyy	12	5519	6	1440	6	4079	294.793312000	0.0908	126875.04	359391.18
xxxx	yyyy	xxxx	yyyy	13	5436	7	3071	6	2365	295.286585000	0.4324	56815.92	43754.37

Table A.5 UDP flows ordered by heavy hitter end point bytes

UDP Conversations

Address A	Port A	Address B	Port B	Packets	Bytes	Packets A->B	Bytes A->B	Packets A<-B	Bytes A<-B	Rel Start	Duration	bps A->B	bps A<-B
xxxx	yyyy	xxxx	yyyy	224	13860	224	13860	0	0	1.592621000	408.4631	271.46	N/A
xxxx	yyyy	xxxx	yyyy	211	13082	211	13082	0	0	1.867859000	409.6521	255.46	N/A
xxxx	yyyy	xxxx	yyyy	196	18320	196	18320	0	0	53.326330000	358.1894	409.17	N/A
xxxx	yyyy	xxxx	yyyy	140	68234	140	68234	0	0	13.073987000	398.3008	1370.50	N/A
xxxx	yyyy	xxxx	yyyy	47	17343	0	0	47	17343	13.074541000	398.3006	N/A	348.34
xxxx	yyyy	xxxx	yyyy	42	15498	0	0	42	15498	17.387569000	374.9915	N/A	330.63
xxxx	yyyy	xxxx	yyyy	36	3312	36	3312	0	0	13.504966000	378.1107	70.07	N/A
xxxx	yyyy	xxxx	yyyy	36	2088	0	0	36	2088	161.783368000	206.5695	N/A	80.86
xxxx	yyyy	xxxx	yyyy	36	5922	0	0	36	5922	17.396969000	348.0261	N/A	136.13
xxxx	yyyy	xxxx	yyyy	36	3312	36	3312	0	0	24.625805000	320.9413	82.56	N/A
xxxx	yyyy	xxxx	yyyy	32	5314	0	0	32	5314	10.819365000	394.9700	N/A	107.63
xxxx	yyyy	xxxx	yyyy	30	5018	0	0	30	5018	32.914875000	370.4624	N/A	108.36
xxxx	yyyy	xxxx	yyyy	30	5286	0	0	30	5286	7.070626000	401.4037	N/A	105.35
xxxx	yyyy	xxxx	yyyy	28	5020	0	0	28	5020	61.595281000	342.7364	N/A	117.17
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	4.265939000	400.2831	126.27	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	4.477096000	400.3840	126.24	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	5.872089000	400.0084	126.36	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	6.801711000	400.0367	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	7.543851000	400.0421	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	8.083161000	400.2283	126.29	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	8.202490000	400.2736	126.27	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	8.420542000	400.0389	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	8.949987000	400.2367	126.29	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	9.768650000	400.2297	126.29	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	9.814400000	400.2656	126.28	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	10.294794000	400.0389	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	10.836523000	400.2669	126.28	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	11.563959000	400.2656	126.28	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	11.701854000	400.0232	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	1.542215000	400.0226	126.35	N/A
xxxx	yyyy	xxxx	yyyy	26	6318	26	6318	0	0	1.811931000	399.9830	126.37	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	4.809094000	399.6452	121.61	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	10.439151000	399.8668	121.54	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	0.000282000	399.6341	121.61	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	13.719239000	384.3963	126.43	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	15.766646000	384.0010	126.56	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	1.986426000	399.7799	121.57	N/A
xxxx	yyyy	xxxx	yyyy	25	6075	25	6075	0	0	2.139889000	399.6688	121.60	N/A
xxxx	yyyy	xxxx	yyyy	24	5832	24	5832	0	0	13.834749000	383.4658	121.67	N/A
xxxx	yyyy	xxxx	yyyy	24	5730	24	5730	0	0	61.586028000	324.5128	141.26	N/A
xxxx	yyyy	xxxx	yyyy	23	3750	0	0	23	3750	0.120452000	364.7750	N/A	82.24
xxxx	yyyy	xxxx	yyyy	21	1932	21	1932	0	0	1.002784000	16.5304	935.00	N/A
xxxx	yyyy	xxxx	yyyy	18	1656	18	1656	0	0	83.998103000	312.7562	42.36	N/A
xxxx	yyyy	xxxx	yyyy	18	1260	18	1260	0	0	295.675066000	9.4102	1071.18	N/A
xxxx	yyyy	xxxx	yyyy	18	1656	18	1656	0	0	318.251467000	34.4897	384.11	N/A
xxxx	yyyy	xxxx	yyyy	14	3366	14	3366	0	0	101.252798000	269.9822	99.74	N/A
xxxx	yyyy	xxxx	yyyy	13	3039	13	3039	0	0	61.558268000	319.0550	76.20	N/A
xxxx	yyyy	xxxx	yyyy	12	1104	12	1104	0	0	167.957557000	128.3790	68.80	N/A
xxxx	yyyy	xxxx	yyyy	12	1104	12	1104	0	0	105.648773000	188.2393	46.92	N/A
xxxx	yyyy	xxxx	yyyy	12	840	12	840	0	0	280.602851000	4.6055	1459.11	N/A
xxxx	yyyy	xxxx	yyyy	10	920	10	920	0	0	65.309028000	259.3985	28.37	N/A
xxxx	yyyy	xxxx	yyyy	10	700	10	700	0	0	295.573730000	9.5113	588.77	N/A
xxxx	yyyy	xxxx	yyyy	10	920	10	920	0	0	342.739936000	12.7680	576.44	N/A
xxxx	yyyy	xxxx	yyyy	8	736	8	736	0	0	42.541685000	360.6525	16.33	N/A
xxxx	yyyy	xxxx	yyyy	8	2266	8	2266	0	0	12.221941000	390.9727	46.37	N/A
xxxx	yyyy	xxxx	yyyy	8	1800	8	1800	0	0	101.251772000	223.0913	64.55	N/A
xxxx	yyyy	xxxx	yyyy	7	1085	7	1085	0	0	42.489239000	359.8852	24.12	N/A
xxxx	yyyy	xxxx	yyyy	7	644	7	644	0	0	46.399822000	324.6865	15.87	N/A
xxxx	yyyy	xxxx	yyyy	7	490	7	490	0	0	280.502317000	4.7058	833.01	N/A
xxxx	yyyy	xxxx	yyyy	6	552	6	552	0	0	55.721919000	3.7547	1176.12	N/A
xxxx	yyyy	xxxx	yyyy	6	552	6	552	0	0	212.730405000	3.7707	1171.14	N/A
xxxx	yyyy	xxxx	yyyy	6	552	6	552	0	0	219.893525000	4.9298	895.78	N/A
xxxx	yyyy	xxxx	yyyy	5	1222	5	1222	0	0	18.958437000	157.2900	62.15	N/A
xxxx	yyyy	xxxx	yyyy	5	1165	5	1165	0	0	342.735775000	12.7731	729.66	N/A
xxxx	yyyy	xxxx	yyyy	4	344	4	344	0	0	314.834780000	86.9902	31.64	N/A
xxxx	yyyy	xxxx	yyyy	4	368	4	368	0	0	113.733295000	271.7616	10.83	N/A
xxxx	yyyy	xxxx	yyyy	4	368	4	368	0	0	141.377150000	151.8045	19.39	N/A
xxxx	yyyy	xxxx	yyyy	4	344	4	344	0	0	38.878838000	85.9882	32.00	N/A
xxxx	yyyy	xxxx	yyyy	3	709	3	709	0	0	26.939436000	355.1166	15.97	N/A

Authors

Sasan Adibi (BS'95, MS'99, MS'05, PhD'10 "expected") is currently involved in the design and implementation of eHealth and mHealth applications, in addition to the next generation wireless and mobile networks, including; Wi-Fi (IEEE 802.11 a/b/g/n) and Cellular (HSPA and LTE). He has an extensive research background mostly in the areas of Quality of Service (QoS) and Security. He is the first author of +25 journal/conference/book chapter/white paper publications and ac co-editor of two books in the areas of 4th Generation Mobile Networks and QoS. He also has strong industry experiences, having worked in numerous high-tech companies, including Nortel Networks, Siemens Canada, WiMAX Forum, and Research In Motion (RIM). He is currently in his last semester at the University of Waterloo.