Optimal Queue Sizing for Traffic Management of Mixed Mobile Traffic in Wireless Network

Zin Win Aye and Myat Thida Mon
University of Computer Studies, Yangon
zinwin2008@gmail.com, myattmon@gmail.com

Abstract

Nowadays, wireless networks which offer various multimedia services to mobile users are demanding Quality of Services (QoS) for these services. Therefore, defining optimal queue size of various mixed applications at an Access Point (AP) not only plays a key role for resource scheduling and QoS of mixed applications and but also is a challenging problem in 802.11 Wireless networks. Due to the fact that queue size impacts on throughput and delay of mixed multimedia applications in congested network, the main objective of this paper is to provide an optimal queue sizing mechanism which maintain high link utilization efficiently and give low queuing delays in wireless networks. The optimal queue sizing method which is based on router buffer sizing methods such as QFind and Rule of Thumb methods is proposed by testing and analyzing the performance of users at an AP with various numbers of queues according to previous proposed papers.

1. Introduction

With the growing increases of mobile operators, the mobile network shares the resource among mobile operators which serve both voice and internet applications to mobile users on the same core network [1]. Although many researches [5] focus on scheduling algorithms such as Proportional Fair, Round Robin in wireless communication environment to improve QoS, one of the most challenging issues in wireless networks is the queue size of APs or Base Stations (BSs) which can directly impact on the unfairness usage of bandwidth of this traffic, flow's achievable throughput and delay. While a very small queue may keep bitrates below the available capacity and packet drops significantly, a large AP queue size can negatively impact a flow's end-to-end delay which is detrimental to VoIP and some network game applications.

As a results of our previous proposed papers [11, 12], there also does not exit specified queue size capable of ensuring both high throughput efficiency and reasonable delay across the range of physical rates and offered loads experienced by modern Wireless Local Area Networks (WLANs). That is, big queue size necessarily carries reduced throughput efficiency and/or excessive queuing delays while very small buffer size causes excessive packet loss and congestion at AP. This, therefore, naturally leads to the consideration of adaptive approaches to queue sizing, which adjust the queues' size in response to depending on the network conditions to ensure both high utilization of the wireless link and avoiding unnecessarily long queuing delays.

In paper [11], a framework with a new scheduling method in WLAN, a two-step traffic scheduling method is proposed to satisfy efficient some QoS parameters: throughput, fairness and delay for mixed applications and to adaptively balances between unicast and multicast applications. The paper [12] proposed priority based scheduling method for resource sharing of voice users in different mobile operators at core node in WLAN. According to these observation in [11, 12], defining optimal queue size at an AP is very important for QoS of mobile traffic in WLANs because many queues at a core or central node in congested networks with various mixed traffic are to hold data right before it is scheduled.

In this paper, there are two different queue models of wireless networks in our proposed papers to give fairly priority assignment for resource allocation in mobile core network. Our proposed network models define four queues and three queues respectively. Therefore, this paper needs to consider queue size of each operator for calculating the priority value of each operator [11] or queue size of each traffic type for calculating the priority value of four mixed traffic types [12] in wireless network. Therefore, an optimal queue sizing method in the system is considered as a main issue for QoE of mixed mobile users to be fairness and to get least delay and high throughput for two types of network topologies**.**

2. Related Works

As the growth of multimedia applications, the main challenge is regarding the enhancement of queue sizing problems for heterogeneous traffics in all wireless networks such as Wi-Fi, WiMAX, LTE and 802.11 networks and so on. Therefore, we analyzed the following related researches concerning buffer sizing methods to guarantee faster delivery and better communication range.

The paper [9] considered the task of sizing buffers for TCP flows in 802.11e WLANs. A number of fundamental new issues arise compared to wired networks. They found that these considerations lead naturally to a requirement for adaptation of buffer sizes in response to changing network conditions. Motivated by these observations they proposed an adaptive buffer sizing algorithm which emulates the classical BDP rule and demonstrated its efficacy via simulations.

The effectiveness of this simple adaptive algorithm can be seen that it maintains high throughput efficiency across the entire range of operating conditions. This is achieved while maintaining the latency approximately constant at around 400ms – the latency rises slightly with the number of uploads due to the over-provisioning to accommodate stochastic fluctuations in service rate. Buffer sizing while rate adaptation is enabled is left as future work, although their proposed algorithm will work. Therefore, it still needs to consider of the possibility of reducing buffer sizes when multiplexing occurs**.**

The paper [10] illustrated the transport layer unfairness problem in the IEEE 802.11 WLANs. They proposed a simple analytical model to calculate the per flow congestion window limit that provides fair TCP access in a wired/wireless scenario where the wireless hop is an 802.11 link. The proposed analysis considered the effects of varying number of uplink and downlink TCP flows, varying Round Trip Times (RTTs) of TCP connections, and the use of delayed TCP Acknowledgment (ACK) mechanism. The comparison with simulation results validates the accuracy of the analytical estimations on the congestion window limit.

When the TCP connections use the congestion window limits calculated from their model, not only fair access can be provisioned, but also the channel can be utilized efficiently in a wide range of scenarios. The analysis could be valuable for deciding on the AP buffer size that provides fair TCP access. The 802.11 vendors may use this method with statistics of TCP connections and WLAN traffic to decide on a good size of AP buffer to optimize the chip size.

Although most buffer sizing algorithms have been proposed to enhance throughput and be low delay for only TCP traffic in wired and wireless networks, there is no still method for fairness problem of mixed traffic in wireless network such as both TCP and UDP or only UDP. Therefore, the main contribution of the paper is to introduce a technique that specifies an optimal queue size level at an AP which has many queues in heavy loaded WLAN which gives the stable fairness index and low delay.

By the result of simulation in Figure 3, 4, and 5, the effect of AP's queue size on fairness problem of mixed traffic type TCP and UDP in [11, 12] proved that the changing buffer size will lead unstable index of heavy loaded traffics. In this paper, a specific range of buffer size is introduced using proposed queue sizing method with the implementation and simulation of various network models in [11, 12] which the best fairness index is achieved with high transmission rate and low delay.

3. Network Topologies

In wireless systems of mixed heavy traffic, how to design the packet scheduling algorithms, resource allocation methods and multiplexing at the packet level to guarantee QoS requirements [3] is not only an important problem, but also a complex problem. The system [11, 12] proposed scheduling and queuing algorithm as in Figure 1 and Figure 2. Considering drawbacks of queuing algorithms [4], the delay is unbalanced depending on the number of packets in each queue because resources are distributed unfairly. According to different network model in Figure 1 and 2, the different queuing size of operators and multiple applications causes the unbalanced queuing delay which affects the throughput, fairness and the delay of mobile users. Therefore, the system presents a queue sizing problem for not only each operator but also mixed mobile traffic.

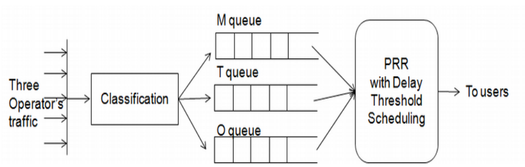

Figure 1. Detail system architecture [11]

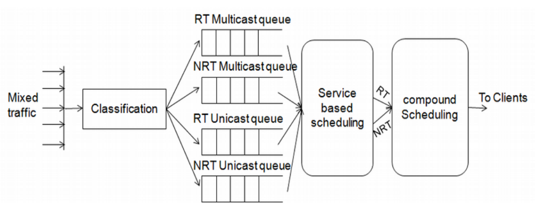

Figure 2. Detail system architecture [12]

Figure 1 presented PRR with Delay Related Scheduling for only voice users among different mobile operators to be high fairness and less delay of users by implementing and testing the system in WLAN. In meanwhile, although the proposed systems utilize scheduling and queuing algorithm which select dynamically the chance of each operator or each traffic type for scheduling users to balance between queue delays, we need to define a unit queue size of all queues on each proposed model to be good fairness among queues because it can adaptively change the number of packets in each queue and its results can impact on the associated priority value for mobile operator or users.

Likewise, Figure 2 proposed scheduling and queuing algorithm for four mixed traffic types of mobile users in wireless network such as real time multicast, real time unicast, non real time multicast and non real time unicast. Firstly, the approach targets the classifying of different mobile users according to their traffic type into suitable queues in core network. Secondly, it schedules each traffic group for defining scheduling time interval of mobile users and schedules each traffic within define time interval to be fairness, to have good throughput, and to be less delay using priority value of multiple queue sizes that is available at AP [12].

4. Evaluation of Proposed System Results

In this paper, it uses NS3 Simulation to test proposed algorithms for four mixed traffic types and three operators in an AP in WLAN, and test fairness level depending on the different queue sizes of two network models using simulation parameters in Table 1. Figure 3 shows the fairness of users who use voice services of three operators at an AP or a core node according to their different queue size. In Figure 4 and 5, it expresses both RT delay fairness and NRT throughput fairness using Jain's fairness index for multicast users and unicast users, respectively.

The simulation outputs consider for four mixed traffic types on both TCP and UDP transport level packet generators in Figure 4 and 5 and for three mobile operators on only UDP packet generator in Figure 3 using parameters in Table 1 and observe that overall fairness of each network model increases and stable for all users at each different queue size because they have different traffic types, network design, number of users, propagation delay depending on the mobile users. Therefore, we need to define the best queue size according to these factors for each network model. Otherwise, the system will face low bandwidth, high packet loss rate and long delay at a congested node in each network topology.

TABLE 1. SIMULATION PARAMETERS

Parameters	Values
Number of AP	1
Channel Capacity	54 Mbps
Packet size	1500 bytes
Channel type	Wireless
Delay threshold	100 ms
Propagation Model	OFDM
Number of users	10
Traffic Types	1 [figure 1], 4[figure 2]
Number of traffics	10[figure 1], 20[figure 2]
Simulation Time	100 s

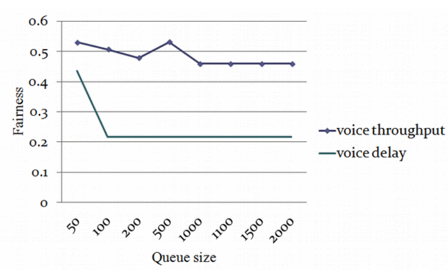

Figure 3. Fairness Results of network model [11]

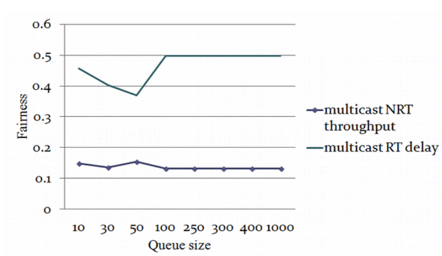

Figure 4. Fairness Results of network model [12]

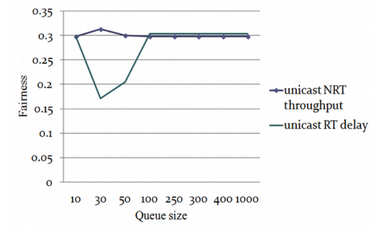

Figure 5. Fairness Results of network model [12]

5. Analysis of Queue Sizing Methods

Despite the importance of queue size to achievable throughput and delay, there is in practice little documentation on AP queue size settings. To investigate queue sizes in wireless APs, a variant of the QFind [6] methodology is applied to infer the access network queue size for the AP or congested nodes. Furthermore, according to the classic rule of thumb the queue size is set to be the product of the bandwidth and the average delay (round trip time or RTT) of the flows utilizing this link: the Bandwidth-Delay Product (BDP) rule [7]. Related sessions and statistical analysis provides details on the queuing delay and throughputs that are then used to compute queue sizes in wired network. In this paper, it is proved that various queue sizes in such a way can fluctuate fairness index of heavy traffics at a core node of different network models based on simulation results, will provide the optimal queue sizes of a congested node in these networks using proposed QS method to satisfy suitable QoS parameters of wireless network.

5.1. QFind

The QFind methodology for measuring wireless AP queue sizes determines the wireless saturation point, measures the baseline delay, induces the saturation rate and measures the delay with queuing, and computes the queue size. QFind is then used to measure the queue sizes targeted for residences. To measure the queue size of wireless access points, a QFind session sends downstream traffic to determine the maximum queuing delay due to the AP queue. The QFind technique has four steps: 1) measure end to end delay on an unloaded network; 2) induce enough network load to fill up the bottleneck queue; 3) while at network load measure the delay and throughput; and 4) use these measurements to infer the queue size of the access. The host AP queue size can then be estimated by the product of the flow throughput and the maximum delay. Let D_t be the total delay (the maximum delay seen by UDP Heartbeat):

$$D_t = D_l + D_q \tag{1}$$

where D_l is the latency (the baseline delay) and D_q is the queuing delay. Therefore:

$$D_q = D_t + D_l \tag{2}$$

Given throughput T (measured at the wireless client), the access link queue size in bytes, q_b, can be computed by:

$$q_b = D_q + T \tag{3}$$

For packet size s (a 1400 byte application payload was used), the queue size in packets, q_p, becomes:

$$q_p = \frac{(D_t + D_l) \times T}{s} \tag{4}$$

5.2. Rule of Thumb

Buffers are currently sized using a rule of thumb which says that each link needs a buffer of size at routers in wired network. This suggests additional engineering and science to determine the best mechanisms for providing AP queue sizes for a variety of traffic types. In this context, the source is the computer initiating the signal and the destination is a remote computer or the system that receives the signal and retransmits it. Having small buffers is attractive as it reduces the amount of memory, required physical space, energy consumption, and price of the router. According to the observation of simulation results in Figure 3, 4, and 5, the main advantage of having small buffers is the reduction in queuing delays and jitter. However, it cause packets drop at small buffers and occurs retransmission and congestion at a center node. Moreover, the average number of hops in the current internet is on a random path. For a single flow with that many hops it is possible to expect several congested links on the path. Thus buffering of several hundreds *ms* at each router would imply very large queuing delays. Using theory, simulation and experiments on a network of real routers, a widely used rule-of-thumb states that each link needs a buffer of size:

$$B = RTT \times C \tag{5}$$

where RTT is the average round-trip time of a flow passing across the link, and C is the data rate of the link, The new rule of thumb with n connections, the number of connections states as follow:

$$B = (RTT \times C) / \sqrt{n} \tag{6}$$

6. Optimal Queue Sizing Method

In this paper, an Optimal Queue Sizing (OQS) method is proposed for a congested node with multiple queues to improve QoS of multimedia traffic in wireless networks. The accuracy of OQS method is validated in two different wireless networks with an AP, where the estimated queue size is compared to the measured queue size from Figure 3, 4, and 5 that gives overall highest fairness level for users. Here, the round trip time (RTT), also called as roundtrip delay is the time required for a single pulse (or) packet to travel from a specific source to a specific destination and back again. That is, it is the twice of propagation delay of each user in the network. And this delay also depends on the distance and transmission speed of users in mobile network model depending on the channel quality of each user. To account the fairness of multiple traffics at an AP according to the impact of buffer size, the proposed queue sizing method is explained in equation 7 to 9.

$$D = \underset{1 \le i \le n}{avg} \left\{ \sum_i RTT_i \right\} \qquad (7)$$

$$EQ = (D.C)/(s.\sqrt{n}) \qquad (8)$$

$$Q = \min\{EQ, Q\max\} \qquad (9)$$

where RTT is the round-trip time of a flow passing across the link, D be the average end to end delay in all link. EQ is the estimated queue size, C is the data rate of the link and s is packet size in bits and n is the number of links in the network. As shown in Figure 3, 4 and 5, Q_{max} is set to be measured queue sizes 1000 and 250. Queue size 1000 is stable fairness index of mixed voice traffic of different mobile operators as in Figure 3 and 250 packets which satisfied stable fairness level of wireless network with mixed mobile traffic in Figure 4 and 5. Finally, the method gives Q which is the optimal queue size for each network model with mixed traffic.

7. Performance Evaluation

In our experiments, we use NS3 Simulation to implement proposed algorithms for 20 traffics of 10 mobile users and 10 mixed voice traffic for three operators of 10 mobile users in an AP in WLAN. The system proposes and tests Optimal Queue Sizing method with different number of traffic using following parameters described in Table 2 for both network models as in Figure 6 and 7 and shows the optimal queue size 322 and 161 for first network model as in Figure 1, and second model in Figure 2 respectively. Furthermore, we compare throughput, delay, and fairness of the measure queue size and estimated queue size using simulation parameters in Table 3 and proposed Optimal Queue Sizing Method for both network models as below in Figure 8 to 13.

TABLE 2. SIMULATION PARAMETERS

Parameters	Values
Number of AP	1
Channel Capacity	54 Mbps
Packet size	1500 bytes
Channel type	Wireless
Link delay[8]	2ms[Figure 1], 1ms[Figure 2]
Access Speed	299792458 m/s
Traffic Types	1[Figure 1], 4[Figure 2]
Simulation Time	10 s

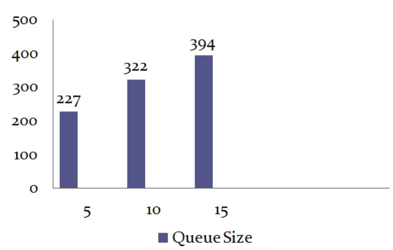

Figure 6. Optimal Queue Sizes with different number of traffic [11]

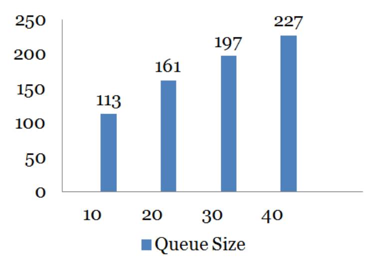

Figure 7. Optimal Queue Sizes with different number of traffic [12]

Table 3. SIMULATION PARAMETERS

Parameters	Values
Number of AP	1
Channel Capacity	54 Mbps
Packet size	1500 bytes
Channel type	Wireless
Delay threshold	100 ms
Propagation Model	OFDM
Number of users	10
Traffic Types	1 [Figure 1], 4[Figure 2]
Number of traffics	10[Figure 1], 20[Figure 2]
Queue Size	1000, 322 [Figure 1] 161, 250 [Figure 2]
Simulation Time	100 s

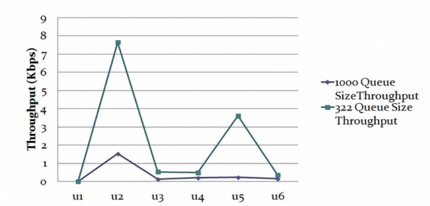

Figure 8. Simulated throughput Results with estimated and optimal queue size [11]

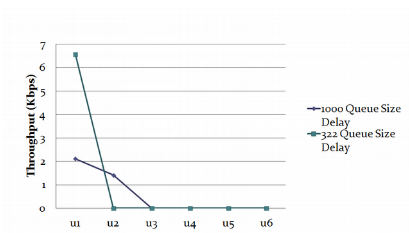

Figure 9. Simulated delay Results with estimated and optimal queue size [11]

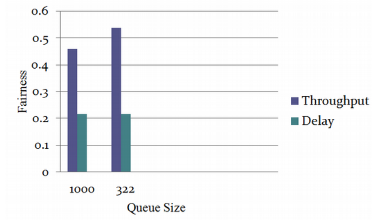

Figure 10. Simulated fairness Results with estimated and optimal queue size [11]

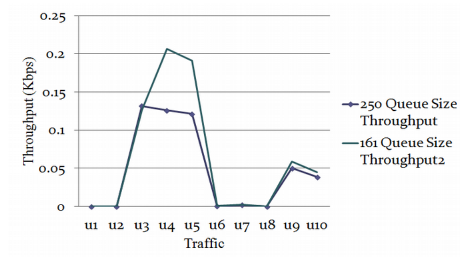

Figure 11. Simulated throughput Results with estimated and optimal queue size [12]

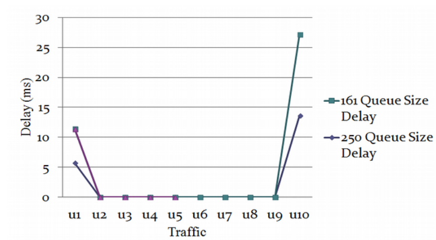

Figure 12. Simulated delay Results with estimated and optimal queue size [12]

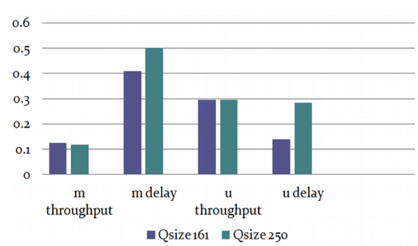

Figure 13. Simulated fairness Results with estimated and optimal queue size [12]

8. Conclusion

The paper proposes the optimal queue sizes of each network model based on the previous proposed framework of traffic management and packet scheduling to increase bandwidth usage and fairness without much delay for mobile users and mobile operators in WLAN. The schemes have likely more performance with a high degree of compatibility with existing scheduling methods for mixed delay sensitive traffic and bandwidth aware traffic and for delay aware traffic such as voice of different mobile operators. Although we implemented this method on only these network topologies at AP in WLAN, it can be applied not only WiMAX but also other wireless and mobile networks. Moreover, it can also be used in any central nodes or many APs with many queues which accepts various multiple traffic in wireless networks although it is implemented at an AP in WLAN as a prototype. As a future work, we will investigate and implement the method to be satisfied with QoS parameters with real wireless environment.

References

[1] Brand, A., Aghvami, H., "Multiple access protocols for mobile communication", http://www.wiley.co.uk or http://www.wiley.com, Swisscom Mobile, Switzerland, King's College London, UK, (2002).

[2] Engin, M. A., Cavusoglu, B., " IEEE communication letter", In: vol 15, no-11, pp. 1234-1237, November (2011).

[3] Manvi, S.S., Venkataram, P., "Mobile agent based online bandwidth allocation scheme for multimedia communication", Electrical Communication Engineering Department, indian Institute of Science, Bangalore-560012 INDIA.

[4] Haas, H. J., "Queuing Methods From FIFO to CB-WFQ/PQ", www.perlhel.at, May (2005).

[5] Tsai, T. Y., Chung, Y. L., Tsai, Z., " Introduction to packet scheduling algorithms for communication network", Institute for Information Industry, Graduate Institute of Communication Engineering, National Taiwan University, and Taipei, Taiwan, R.O.C, http://www.intechopen.com.

[6] Mark Claypool, Robert Kinicki, Mingzhe Li, James Nichols, and Huahui Wu., " Inferring Queue Sizes in Access Networks by Active Measurement", CS Department at Worcester Polytechnic Institute, Worcester, MA, 01609, USA, 2003.

[7] Guido Appenzeller Isaac Keslassy Nick McKeown., "Sizing Router Buffer", Stanford University Stanford University Stanford University, SIGCOMM'04, Aug. 30–Sept. 3, 2004 , ACM 1-58113-862-8/04/0008, Portland, Oregon, USA.

[8] "Real-Time Traffic over Wireless LAN Solution Reference Network", http://www.cisco.com, November 11, 2013, Americas Headquarters, Cisco Systems, Inc.,170 West Tasman Drive, San Jose, CA 95134-1706, USA.

[9] Tianji Li, Douglas J., " Buffer Sizing for TCP Flows in 802.11e WLANs", Leith Hamilton Institute, National University of Ireland Maynooth, Ireland, supported by Science Foundation Ireland Grant 03/IN3/I396.

[10] Reza Adinehnia and Mohamed Othman., "Direction-based TCP Fairness Enhancement in High speed IEEE 802.11n", Department of Communication Technology and Network, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor D. E., Malaysia, 2014 ACADEMY PUBLISHER doi:10.4304/jnw.9.8.2013-2020, JOURNAL OF NETWORKS, VOL. 9, NO. 8, AUGUST 2014.

[11] Z. W. Aye and M. T. Mon., "Priority based Scheduling for Resource Allocation of Mobile Operators in Mobile Network", in Proceedings of International Conference on Science, Technology, Engineering and Management (ICSTEM), Thailand, 2015.

[12] Z. W. Aye and M. T. Mon., "Service based Packets Scheduling for QoS of Mixed Mobile Traffic in Wireless Network", 6th International Conference, ICSI and the Second BRICS Congress, CCI, Beijing, China, Proceedings, Part II, Lecture Notes in Computer Science 2015.