A Novel Scheduling and Queue Management Scheme for Multi-band Mobile Routers

Md. Shohrab Hossain, Husnu S. Narman and Mohammed Atiquzzaman
School of Computer Science, University of Oklahoma, Norman, OK 73019
Email: {shohrab, husnu, atiq}@ou.edu

Abstract—Recent trend in Internet usage have seen large amount of multimedia data due to increasingly large number of mobile users. To facilitate higher bandwidth, modern mobile routers are capable of supporting simultaneous multi-band, leading to less interference, higher capacity and better reliability. However, there exists no previous work that attempt to maximize utilization of available bandwidth through the sharing of traffic classes among different frequency bands of the mobile router. In this paper, we have proposed a novel scheduling algorithm for multi-band mobile routers which transmits different classes of traffic through different frequency bands to achieve improved performance. We have developed an analytical model to perform queuing analysis of the proposed multi-band system and derived various performance metrics, validated by extensive simulations. Our results show that the proposed architecture can ensure maximum possible utilization through sharing of capacities among the bands. Our proposed scheduling algorithm and related analytical model can help network engineers build next generation mobile routers with higher throughput and utilization ensuring least packet loss for different classes of traffic.

Index Terms—Analytical modeling, scheduling algorithm, queuing system, real-time traffic, next generation mobile routers.

I. INTRODUCTION

In recent years, there have been explosive growth of mobile users accessing large multimedia files (such as, high definition audio, video, images, etc.) over the Internet. Therefore, the bandwidth demand for mobile Internet access is increasing exponentially [1]. To satisfy such a higher bandwidth requirement, today wireless routers are available commercially with simultaneous multi-band support of 2.4 and 5 GHz. Future IEEE 802.11ad (WiGig) tri-band enabled devices, operating in the 2.4, 5 and 60 GHz bands, are expected to deliver data transfer rates up to 7 Gbps [2]. The benefit of using multi-band router is to have less interference, higher capacity and better reliability. Exploitation of rarely-used frequency bands in wireless networks reduces interference in heavily-used frequency band, e.g., 2.4 GHz, thereby increasing total capacity of the wireless network.

Current simultaneous multi-band MRs make use of two different bands (2.4 GHz and 5GHz) for different types of devices in a home network. However, they do not attempt to exploit the under-utilized frequency band while other one is flooded with data. Moreover, there are different classes of traffic and some of the traffic type (such as, real-time) has some strict delay constraints associated with it; some other signaling traffic (required for mobility management) are crucial for maintaining Internet connectivity of the mobile users. Therefore, it is essential to propose appropriate scheduling and queue management scheme for the multi-class traffic to ensure maximum possible utilization of the system resources in multi-band mobile routers [3]. The aim of this work is to ensure maximize utilization of multi-band mobile routers through band sharing.

There have been several research works [1]–[7] reported in the literature that attempt to extend current single band technology through the use of multiple frequency bands, leading to increased bandwidth while reducing interference. Even though multi band usage has been widely investigated in cell networks [5], [6], it is relatively new concept in wireless networks. Verma and Lee [7] explain possible Wi-Fi architecture with multiple physical and link layers to support multiple frequency bands simultaneously. Singh et al. [1] proposed a method to assign end-devices different bands based on their distances from the access router. In [2], [4], authors proposed the use of 60 GHz frequency band (having low range) to attain faster data transfer rate in wireless networks. However, none of these works [1]–[7] propose any scheduling algorithm for multi-band system considering multi-class traffic; neither do they perform any queueing analysis to measure difference performance metrics.

To the best our knowledge, there has been no earlier works on scheduling and queue management for multi-band mobile routers that attempts to maximize utilization of available bands. Moreover, no previous work exists that proposed the sharing of multiple bands to transmit different classes of traffic. This is a novel work that aims at attaining maximum possible band utilization and proposes a scheduling algorithm which exploits band-sharing.

The objective of this work is to analyze the performance of multi-band mobile routers while ensuring maximum possible utilization through sharing of bands among different classes of traffic.

The contributions of this work are: (i) proposing a band-sharing router architecture and a novel scheduling algorithm to ensures maximum possible utilization of the system, (ii) developing an analytical model to evaluate the performance (utilization of bands, average class occupancy, packet drop rate, average delay, and throughput) of the proposed multi-band system, and (iii) validating our analytical model by
extensive simulations. Our proposed algorithm considers multi-class Internet traffic and schedules them through alternate under-utilized frequency bands, thereby reducing packet loss and delay. Results show the packet drop rate and throughput are significantly improved in proposed band-sharing architecture of the mobile router. Moreover, the simulation results validates our analytical model. Our proposed scheme and related analysis will help network engineers build next generation mobile routers with higher throughput and utilization, ensuring least packet loss of different classes of traffic.

The rest of the paper is organized as follows. In Section II, we explain the typical architecture of multi-band mobile routers, followed by the proposed architecture in Section III. Section IV presents the analytical model to derive different performance metrics of the proposed architecture. In Section V, we present the simulation results that validates our analytical model; we also compare the proposed and typical architecture. Finally, Section VI has the concluding remarks.

II. TYPICAL MULTI-BAND ROUTER ARCHITECTURE

Commercial (simultaneous) multi-band mobile routers available today makes use of two different bands (2.4 GHz and 5 GHz) for different types of devices in a home network. Laptops may connect to 2.4 GHz network while WiFi-enabled TV, gaming devices may connect to 5 GHz network. This reduces interference with the heavily-used 2.4 GHz network (as cordless phones, microwave oven use similar band). In addition, video streaming can be done through the high frequency band. Future IEEE 802.11ad (WiGig) tri-band enabled devices, operating in 2.4, 5 and 60 GHz bands, are expected to deliver data at a much higher rate (up to 7 Gbps) [2].

The main principle of today’s simultaneous multi-band MR is the non-sharing of bands among different flows of traffic. Moreover, some of the devices today (such as, IPTV) mostly deals with real-time traffic. Based on this fact, we have assumed that each of the band of typical simultaneous multi-band MR only deals with one type of traffic. This might be a slight deviation from the real MR used today. However, we have assumed this to compare our proposed architecture with typical simultaneous multi-band MR.

III. PROPOSED MULTI-BAND ROUTER ARCHITECTURE

In this section, we explain the proposed architecture of multi-band MRs that promotes sharing of bands to maximize system utilization. We have considered three different queues (shown in Fig. 2), each of which corresponds to a frequency band of a simultaneous tri-band Mobile Router. As in Fig. 1, we consider three classes of traffic and each queue is designated for each class of traffic. However, unlike the typical architecture, in this proposed architecture (see Fig. 2) traffic of one class can flow through other queues provided the other queues have empty slots, thereby ensuring better utilization of buffer spaces available. For example, if the B-queue has some empty spaces available and a bursty RT traffic comes in, the overflowed RT traffic can be queued in the B-queue and subsequently served (or sent) through the B-server (transmitter).

A. Time and space priority

The time and space priority for the three queues of the proposed architecture are explained in Figs. 3, 4 and 5. For B-queue, BU packets have the highest priority; RT and NRT packets have dynamic priority based on arrival rates (see Eqns. (4) and (5)). Regarding space priority, BU packets are queued in front of B-queue and if there are empty spaces available, other types (RT and NRT) can be accommodated as shown in Fig. 3.

R-queue can have only RT and NRT packets as shown in Fig. 4. RT traffic has higher priority over NRT traffic.
Therefore, R-queue can have NRT packets only if RT packets cannot fill the R-queue at any instant and there are NRT packets overflowed from the B-queue.

Finally, Fig. 5 shows the N-queue which is designated for NRT traffic. However, if there are empty spaces available in this queue, overflowed RT traffic out of B-queue can be enqueued in N-queue (see Fig. 2).

### B. Scheduling algorithm

We have considered the following two crucial factors to ensure improved performance of the multi-band MR:

- The unused buffer space of one queue (or band) can be used for other traffic types, thereby reducing the idle time of the system.
- Priorities of different traffic classes are also considered while selecting a particular type of packet over others. Priority has an inverse relation with the arrival rate.

Scheduling policies are explained as follows:

- Different class of traffic are first tried to be sent through the designated frequency band.
- If N-queue and/or R-queue are overflowed, corresponding traffic is forwarded to B-queue.
- If B-queue does not have enough empty slots to handle these overflowed NRT and RT packets, they race for slots in B-queue based on priority (see Eqns. (4) and (5)).
- If overflowed RT packets cannot be accommodated in B-queue, they attempt to be queued in N-queue, depending on the available empty space in N-queue.

- If the R-type packets cannot be even accommodated in N-queue, they are dropped from the system.
- Similar policy is enforced while dealing with NRT packets in B-queue and then in R-queue.

### IV. ANALYTICAL MODEL

In this section, we present our analytical model to derive various performance metrics of the proposed multi-band mobile router architecture.

#### A. Assumptions

To make the model analytically tractable, the following assumptions have been made.

- Packet arrival follows Poisson distribution.
- Type of queue discipline used in the analysis is FIFO with non-preemptive priority among various traffic classes.

#### B. Notations

The notations used in the analysis are listed below. To simplify our notation, we use $T$ as the common notation for different traffic class types, and $T \in \{ B, N, R \}$.

- $N_T$ Queue size of $T$-queue in the MR.
- $\alpha_T$ Total packet arrival rate at $T$-queue of $i$-th MN.
- $\mu_T$ Service rate at $T$-queue of $i$-th MN.
- $\sigma_{T_{\text{eq}}}$ Priority of class-$T$ traffic in B-queue.
- $P_{d_{TXQ}}$ $T$-type packet drop probability in X-queue, where $X \in \{ B, N, R \}$.
- $E(D_T)$ Average queuing delay of class $T$ packets.
- $E(n_{T})$ Average queue occupancy of class $T$ packets.
- $P_{d_{\text{sys}}}$ Final packet drop probability of class $T$ packets.

#### C. Total arrival rates in each queue

For queuing analysis of the proposed system, we need to determine the total arrival rate of all class of traffic in each queue. In general, overflow in a queue can happen when the arrival rate is larger than the buffer space or when the service rate is smaller than the arrival rate and there is no buffer space left, as specified by the following conditions:

$$\alpha_T > N_T \quad \text{or} \quad \mu_T < \alpha_T \quad (1)$$

1) $B$-queue: For $B$-queue, number of packets overflowed from N-queue and R-queue goes to B-queue. Thus, the arrival rates of N-type and R-type packets to the B-queue (denoted by $\alpha_N$ and $\alpha_R$) can be obtained as follows:

$$\alpha_N = \alpha_N P_{d_{N_{\text{eq}}}} \quad (2)$$

$$\alpha_R = \alpha_R P_{d_{R_{\text{eq}}}} \quad (3)$$

where $P_{d_{N_{\text{eq}}}}$ and $P_{d_{R_{\text{eq}}}}$ are packet drop probabilities of N-type packets in N-queue and R-type packets in R-queue.

Priorities of different classes are taken into account while allowing traffic into B-queue. Priority of B packets in B-queue is $\sigma_{B_{\text{eq}}} = 1$. Priorities of other classes of traffic in B-queue are measured as follows:

$$\sigma_{B_{\text{eq}}} = \frac{\alpha_R}{\alpha_B + \alpha_R + \alpha_N} \quad (4)$$
The total arrival of N-queue is the sum of two arrival rates $\alpha_N$ and $\alpha'_N$ (see Eqn. (7)); the former has the higher priority than the latter. Therefore, following a similar approach as in Eqn. (11), we can compute $P_{dRNQ}$. Then we can follow similar approach in Eqn. (12) to compute $P_{dRNs}$. The computation of N-type packet drop probability follows similar steps as followed for R-type packets. Therefore, $P_{dN}^{Nys} = P_{dN_{BQ}}$.

### E. Average queue length

Each queues behaves as M/M/1/N queue. Therefore estimated queue length can be obtained as follows:

$$E(n_T) = \begin{cases} \rho_T & \text{if } \rho_T < 1 \\ \frac{N_T}{2} (1 - \rho_T^N) & \text{if } \rho_T = 1 \end{cases}$$

(14)

The average queue occupancy of R-type depends on the queue occupancy of R-packets in R-queue, B-queue and N-queue. This is computed as follows:

$$E(n_R^{sys}) = E(n_B^{RQ}) + E(n_N^{BQ}) + E(n_N^{NQ}) = (\rho_B RQ) + (E(n_B^{BQ}) - E(n_B^{BQ})) + (E(n_N^{NQ}) - E(n_N^{NQ}))$$

(15)

For computing $E(n_R^{RQ})$, we need to put $N_T = N_R$, $\rho_T = \rho_R = \rho_R/\mu_R$ in Eqn. (14). Similar approach can be used for the rest of the terms in Eqn. (15).

Since B-type packets are only queued in B-queue, therefore,

$$E(n_B^{RQ}) = E(n_B^{BQ})$$

(16)

To compute average queue occupancy of N-type packets in the system, a similar approach as in Eqn. (15) can be used:

$$E(n_N^{sys}) = E(n_N^{BQ}) + E(n_N^{NQ})$$

(17)

### F. Throughput

The throughput of T class of traffic can be obtained as follows:

$$\gamma_T^{sys} = (1 - P^{sys}_{dT}) \alpha_T$$

(18)

### G. Average packet delay

Thus, the average packet delay of T class is given by,

$$E(D_T^{sys}) = \frac{E(n_T^{sys})}{(1 - P^{sys}_{dT}) \alpha_T}$$

(19)

### V. RESULTS

We have written discrete event simulation programs in MATLAB environment by taking into account the assumptions and scheduling policies mentioned in Sections II and III. We have followed M/M/3/N [8] procedures for the implementation of simulation programs. We have kept equal buffer length (50 packets) for each queue. Buffer lengths are kept small [10] as in real routers to decrease packet delay. RT and NRT packets are assumed to be 512 bytes [2], [11] whereas the BU packets are assumed to be 64 bytes. The service rates of the B, N and R-queues are kept 27, 75 and 132 packets/sec which is proportional to service rates of multi-band routers [2]. We ran each simulation for 20 trials having different traffic class arrival rates as follows:
\[ \lambda_B(i) = \{ i \}, \lambda_N(i) = \{ 3i \}, \lambda_R(i) = \{ 18i \} \]

where \( i = 1, 2, 3, \ldots, 20 \). We have run simulations with increased arrival rates of all types of traffic to observe the impact of heavy traffic on the multi-band system. The arrival rate of B-queue and N-queue are increased slowly in each trial whereas the RT traffic arrival rate are increased at a much higher rate. This will saturate the R-queue and we explain the impact of this overflow on different performance parameters of our proposed system and typical existing system.

A. Validation of analytical model

In this subsection, we show the simulation results for proposed architecture and compare them with those produced by the analytical expressions derived in Section IV to validate our model.

1) Average class occupancy: Fig. 6 shows the average class occupancy of the proposed multi-band system obtained through simulations and analytical model. The simulation and analytical results are very close to each other. The class occupancy of N-class and B-class are very low as their service rate is higher than arrival rates. However, this is not the case for R-class where the arrival rate is higher than R-queue service rate; hence, the excessive RT packets are enqueued in other two queues, thereby increasing the average occupancy of R-class.

2) Class drop rate: Fig. 7 shows the class drop rate for the proposed multi-band system. Again the simulation results closely match the analytical one. Drop rates of B and N-class are low due to low arrival rate. However, the R-class drop rate rises as we use very high arrival rates in subsequent trials.

3) Class throughput: Fig. 8 shows the class throughput for the proposed multi-band system. Again the simulation results closely match the analytical one.

Thus, it is evident from Figs. 6, 7, and 8 that the analytical and simulation results are very close to each other, thereby validating the analytical model.

B. Comparison between proposed and typical architectures

In this subsection, we present results comparing the performance results of our proposed (band-sharing) architecture and typical (non-sharing) architecture.

1) Utilization: Utilization is a performance measurement that indicates how efficiently bands are used and whether there is any unused capacity of the system. Fig. 9 shows band utilizations for proposed (shared) and typical (non-shared) architectures of multi-band routers. When packet arrival rates are low (in trials #1 through 7) compared to the capacity of each queue, all the queues have somewhat similar and low utilization for both the architectures. However, for trials #8 through 20, the utilizations of B-queue and N-queue are much higher for proposed architecture than for typical one. This is because increased number of RT packets are dropped in typical architecture (see Fig. 11) whereas in proposed one, they are accommodated in B-queue and N-queue, thereby improving their utilizations and maximizing system performance.

2) Average packet delay: Fig. 10 shows the average packet delay of each class of traffic for proposed and typical architecture. The delay for RT traffic (for trials #8 through 20) in proposed architecture is higher than the typical one. This is because excessive RT packets are immediately dropped from the system in typical architecture and these lost packets do not come into account in delay calculations. On the contrary, in proposed architecture overflowed RT packets get chances to be enqueued in B and N-queue before being dropped. RT packets are second priority packets in N-queue and B-queue and they have to wait for NRT and BU packets, respectively before being scheduled for service. Hence, it increases the delay of RT packets.

3) Packet drop probability: Fig. 11 shows packet drop probability of each class for proposed and typical architecture. For both the approaches, drop rate of BU and NRT are low and similar since arrival rates are lower than the service rates. When the arrival rates for RT increases (in trials #7 through 20), RT packet drop rate gradually goes up for typical (non-shared) architecture. However, proposed architecture does not allow RT traffic to drop as long as the excessive RT packets can be queued in the empty buffer spaces at B-queue and N-queue. Therefore, RT packet drop rate is much lower for proposed architecture than the typical one.

4) Throughput: Fig. 12 shows throughput of each class for proposed and typical architectures. The throughput of NRT and BU class are increased with the increase of their arrival
For the typical architecture, the throughput is saturated at $\mu R (= 132 \text{ pkts/sec})$ when the RT arrival rate reaches this value. However, the RT class throughput for the proposed architecture decreases slowly due to the impact of increased arrival rates of other queues (B and N-queue) that results in less available space for overflowed RT packets.

C. Summary of results

Our simulation results closely matches the analytical ones (see Figs. 6, 7, and 8), thereby validating our analytical model. Moreover, we find that the proposed architecture attempts to maximize system utilization (Fig. 9) through band sharing. This affects on the average queue occupancy and delay (Fig. 10) of RT traffic. However, the packet drop and throughput (Figs. 11 and 12) are significantly improved in the proposed band-sharing architecture.

VI. CONCLUSION

In this paper, we have proposed a novel scheduling algorithm for multi-band mobile routers that exploits band sharing. We have developed an analytical model to perform queuing analysis of the proposed multi-band system and derived various performance metrics that have been validated by extensive simulations. Our results show that the proposed architecture can ensure maximum possible utilization through the sharing of capacities among the bands. Our proposed scheduling algorithm and related analytical model can help network engineers build next generation mobile routers with higher throughput and utilization ensuring least packet loss for different classes of traffic.

REFERENCES