A Cross-Layer Metric for Application-Constrained MAC-Aware Capacity Optimization

Bahador Amiri  
Department of Electrical Engineering  
University of California  
Santa Cruz, CA 95064  
Email: bamiri@soe.ucsc.edu

Hamid R. Sadjadpour  
Department of Electrical Engineering  
University of California  
Santa Cruz, CA 95064  
Email: hamid@soe.ucsc.edu

Abstract—This paper introduces a new cross-layer design metric called Application-Constrained MAC-Aware Capacity (ACMAC). ACMAC considers outage capacity, MAC layer and application layer Quality-of-Service (QoS) requirements to derive the lower and upper bounds of the transmission rate. Theoretical derivation of this metric demonstrates the importance of considering different layers of the network. Further, we show how to use ACMAC as a cross layer optimization criterion to generate frameworks for selecting different parameters such as transmit rate, channel or route selection in wireless networks. Our simulation results show that ACMAC is very accurate in predicting network performance such as throughput while satisfying QoS parameters like Packet Error rate (PER).

Keywords—Cross-layer Optimization; Quality-of-Service; Outage Capacity

I. INTRODUCTION

Wireless technology has become an integral part of today’s life covering many applications such as phone calls, data transfer, web browsing, etc. As a result, innovative technologies have been developed and implemented to increase throughput and reliability of different wireless networks including Wireless Local Area Networks (WLAN). Some of the proposed methods have been adopted in IEEE 802.11 standards and resulted in significant performance improvements in the latest standards such as IEEE 802.11ac. These new standard features provide variety of ways for optimizing WLAN performance with a focus on physical and MAC layer improvements. They have introduced improvements such as higher bandwidth (up to 160MHz), higher order modulation (up to 256 QAM), advanced Multiple-Input Multiple-Output (MIMO) techniques including beamforming and Multi-User MIMO (MU-MIMO) capability and Low-Density Parity Check (LDPC) code in the physical layer. With these improvements, WLAN systems are capable of providing data rates faster than 1 Gb/s at short distance [1].

However, a throughput close to this nominal data rate is not achievable in most practical conditions. In addition to networking overheads, the wireless link quality and utilization significantly impacts the achievable throughput of these systems. Due to shared nature of wireless medium, different neighboring systems which operate in the same frequency will share the channel. As a result, WLAN users may experience lower performance in dense wireless environments. This signifies the importance of considering medium utilization while calculating the real-time achievable capacity of wireless networks. One of our contributions in this paper is to provide MAC-awareness by formulating the impact of medium utilization into the system performance.

In [2], the importance of cross layer design for efficient network resource allocation is stated. Further, the importance of modeling network performance and understanding properties of service types over wireless networks along with the need for adaptive control in wireless networks to ensure proper scheduling and utilization of capacity are emphasized. Our work is aimed to provide a cross-layer performance model for analyzing and control of the wireless networks based on different service types.

WLANs were traditionally used for non sensitive applications such as web browsing. Recent technology advancements have brought an increase demand for usage of delay and throughput sensitive applications. Use of mobile devices in WLAN for Over-The-Top (OTT) voice and video services are growing rapidly. Meanwhile, many service providers are using WLAN to distribute live video to multiple clients inside the home. The variation of WLAN applications with their unique Quality-of-Service (QoS) requirements, signifies the importance of considering constraints which come along with each application. One of the contributions of this paper is considering application throughput, packet error rate and delay requirements while evaluating achievable network capacity and dynamically optimizing network parameters based on each application.

Consequently, we combine application and medium constraints into the definition of outage capacity. This provides an Application-Constrained MAC-Aware Capacity (ACMAC) formulation which expands from physical layer to application layer. ACMAC formulation provides a generalized network optimization framework for different purposes such as adaptive rate control, channel selection and network routing.

The outline of the paper is as follows. Section II provides an overview of the related work. Section III describes ACMAC contributing factors from different layers of network. Section IV presents details of our optimization formulation and derivation of its constraints. Section V shows our simulations results and comparison of our metric performance with some other
well known metrics. We conclude the paper in section VI.

II. RELATED WORK

A cross layer rate adaptation method which involves high-level interaction between the MAC and physical layers is introduced in [3]. In this work, two protocols for rate adaptation, Loss-triggered and SNR-triggered are studied. The Loss-triggered protocol is performed by transmitter alone based on received acknowledgments. The SNR-triggered protocol requires receiver to include estimated SNR in the Clear To Send (CTS) message. Results show that SNR-based protocol is able to achieve higher throughput when compared to loss-triggered protocol in many areas including mobility and interference-prone environments.

Another Phy-MAC rate adaptation protocol called SoftRate is proposed in [4]. To avoid training required for SNR estimation, this protocol uses confidence information calculated by the physical layer to estimate the prevailing channel bit error rate (BER). Since the bit-error rate that is calculated may include the interference errors caused by packet collisions, the SoftRate receiver uses a heuristic approach to separate out errors caused by interferers.

Another parameter which can be optimized based on cross layer information is operating frequency or channel for the wireless system. For example, there are several different operating frequencies/channels considered in IEEE 802.11 and each network is allowed to select its desired frequency for operation. Since, operating frequency can impact network performance significantly, it has been studied by many researchers. One of these methods is Non-Utilized Outage Capacity, NUOC [5] which is a cross layer outage capacity metric for wireless channel selection. NUOC considers phy and MAC layer information.

Researchers have also proposed several cross-layer protocols for route selection in ad hoc networks. Some of these protocols are between physical and network layers. Authors in [6] have proposed Signal strength assessment Based Route Selection for OLSR (SBRs-OLSR) where the link connectivity is based on SNR measurement from Phy layer. SNR and received power are considered in [7] and outage capacity is considered in [8] for route selection purposes.

There are other cross layer routing methods which are between MAC and network layers [9], [10]. In [9] a routing protocol based on movement prediction is proposed which improves the routing process by taking MAC layer movement information into consideration. The authors in [10] use a routing metric which is based on hop count and transmission counts at MAC layer.

Cross layer design based on Quality-of-Service (QoS) is considered in literature. Authors in [11] have provided a method to employ joint source coding (application-layer), channel coding and rate adaptation (physical layer). A joint application-layer scheduling and MAC-layer retransmission strategy is developed in [12]. In this work, application and MAC layers jointly decide the optimal packet size and retransmission limits. In [13], the authors have utilized a data partitioning technique at the application layer and QoS mapping technique at the MAC layer of the 802.11e network. Another cross layer design is proposed in [14] which maps video packets at the application layer to the appropriate MAC layer access categories in IEEE 802.11e.

There are other areas of cross layer research works like transmit power optimization which is beyond the scope of this paper. There are survey papers in cross layer design which can be used for further study in this area [15], [16]. The major difference between our research and prior works is that we provide a multi-purpose cross-layer optimization and evaluation framework. This framework can be used for many different purposes including rate adaptation, frequency selection, routing and QoS optimization.

III. ACMAC CROSS LAYER FACTORS

In this section, we provide details on the design of our cross-layer approach. These factors are from physical, MAC and application layers.

A. Physical Layer

We consider Bandwidth (BW) and Signal to Interference plus Noise Ratio (SINR) as two factors which can provide good estimation of channel quality. These two factors are used in channel capacity computation. Unlike SNR which can vary significantly over time, bandwidth is usually a static value in wireless systems.

There are other factors like Received Signal Strength Indicator (RSSI) which are used by other researchers as indicator for link quality. RSSI which is measured by receiver is a function of transmit power and channel loss. SINR by definition includes RSSI. SINR from node $i$ to node $j$ can be described as

$$SINR_{ij} = \frac{P_i}{N + \sum_{k \in \mathcal{N}(j)} 1 + L_{kj}},$$

where $P_i$ and $P_k$ are transmit power strength at node $i$ and $k$, $L_{ij}$ is the channel loss from node $i$ to $j$, $\eta(j)$ is the set of neighbors for node $j$, $L_k$ is the channel loss from interfering node $k$ to node $j$ and $N$ is noise. In WLAN systems, interference is usually a more significant factor than noise, therefore for the remaining of this paper we use interference as a term representing both noise and interference in SINR calculation. In section IV, we show how SINR is used in ACMAC formulation.

B. MAC Layer

Due to shared nature of wireless medium, a multiple access protocol is required to make sure different devices are able to reliably access the medium for transmission. Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) is the MAC protocol used in IEEE 802.11 standard. With CSMA/CA, before transmission, each device listens to the medium for a predetermined time interval. If the medium is idle during this interval, the device is allowed to start transmission. If the medium is busy, the device has to defer
transmission and wait for the medium to become available. Collision avoidance is provided by a random back-off time before transmission. Since collision penalty in wireless systems is significant, the random back-off is designed to minimize the collision possibility.

Carrier sensing part of CSMA is carried by measuring the energy on the channel and comparing it against a predetermined threshold. If the energy was higher than the threshold, the medium is considered busy; otherwise, the medium is considered idle and available for transmission. A time unit for each device can be divided into four portions: 1) Tx time: the time spent by the device to transmit packets, 2) Rx time: the time spent by the device to receive packets, 3) Idle time: the time the medium is idle and not used by device or any interfering traffic [8].

Channel availability is inversely proportional with the interference in the close proximity, i.e., (any interfering traffic [8]. The time the medium is used by interfering devices and 4) Idle time: the time the medium is idle and not used by device or any interfering traffic [8].

As explained earlier, WLAN are used for many different applications such as voice, video and data transfer. Each of these applications have their own Quality-of-Service (QoS) requirements. Three most important QoS parameters for many applications are throughput, latency and Packet Error Rate (PER). It is important to mention that application requirements are defined based on the end-to-end network performance, while in this paper we use these requirements for the wireless portion of the network. Limitations in other parts of the network are beyond the scope of this paper.

In order to design a system with good user experience, these QoS parameters should be considered closely based on network application. Many of the previous research works have been either focused on optimizing lower network layers without considering specific application QoS requirements [3]–[10] or have proposed QoS aware methods for a particular application like video [13], [14].

Our metric, ACMAC, is an application aware metric. It means that it looks at current network application requirements and defines constraints on the data transmission rate accordingly to satisfy these requirements. As a result, system will dynamically adapt itself to any network application to provide customized QoS for that application. Moreover, network will achieve different capacity depending on specific application constraints. This scheme is designed guarantee high QoS for variety of applications. As a result, each application will have its own unique capacity optimization criterion which is dynamically determined.

ACMAC is also a multi-purpose metric. It means that this metric is not designed for optimization of a particular parameter. Similar to capacity which can be used as a metric for many different types of protocol designs, ACMAC can also be used for optimization of wide variety of parameters such as rate selection, channel selection or routing.

IV. CROSS-LAYER DERIVATION

As stated earlier, our scheme considers BW and SINR from phy layer, channel utilization from MAC layer and QoS requirements (i.e. latency, packet loss and throughput) from application layer. In this section, we show derivation of a multi-purpose network optimization metric which considers all above mentioned factors.

We consider a Rayleigh fading channel modeled as a complex Gaussian random variable with zero mean and variance $\sigma^2_m$ per complex dimension. An additive white Gaussian noise with variance $N_0/2$ per complex dimension is added to the received signal. With a finite bandwidth of $B$ Hz and transmit power of $P_T$, the average received SNR $\gamma_m$ is [17]

$$\gamma_m = \frac{P_T}{B N_0}$$

(2)

The outage probability for a Rayleigh fading channel is given by

$$P_{out} = P(\gamma < \gamma_{min}) = 1 - \exp\left(-\frac{\gamma_{min}}{\gamma}\right),$$

(3)

where $\gamma_{min}$ is the minimum SNR required for supporting a certain data rate. Outage probability is an important parameter and has been considered as optimization criteria [18], [19]. For an application with outage probability requirement of $P$, one constraint which needs to be satisfied is $P_{out} \leq P$. By utilizing equation 3, we arrive at

$$1 - \exp\left(-\frac{\gamma_{min}}{\gamma}\right) < P$$

$$\gamma_{min} < -\gamma \ln(1 - P)$$

(4)

Considering capacity definition of $R = B \times \log_2(1 + \gamma_{min})$, we have

$$\gamma_{min} = 2^{R/B} - 1.$$ 

(5)

Combining equations (4) and (5) will result in

$$R < B \times \log_2 (1 - \gamma \times \ln(1 - P)).$$

(6)

This equation defines an upper bound for the transmit rate where $B$, $\gamma$ and $P$ are bandwidth, SINR and maximum acceptable outage probability respectively. In practice, there is a big difference between $P$ and application packet loss requirement, $P_L$. $P$ is the probability of transmission failure for a single attempt while application packet loss occurs only after several retransmission failure at lower layers.

In most deployed systems, the maximum number of retransmissions is static and is a design decision. What we suggest here is that this number should also be determined based on application. Different applications have different latency requirement. For example, packets generated by a voice
application have much more strict latency requirements than packets corresponding to file transfer. Since each retransmission will take certain time, \( T \), the number of retransmissions for voice call packets should be much less than file transfer to satisfy strict latency requirement. We define the maximum number of retransmissions as \( N = \lceil L/T \rceil \) where \( L \) is latency requirement specific to an application. In systems, where static number of retransmissions is used, \( N \) can be replaced by this static number. The relation between \( P \) and \( P_L \) is defined as \[8\].

\[ R \leq R_{MAX} = B \times \log_2 \left( 1 - \gamma \times \ln(1 - \sqrt{\frac{P_L}{B}}) \right) \tag{7} \]

This equation provides a cross layer upper bound of transmission rate \( R_{MAX} \) based on bandwidth \( (B) \), SINR \( (\gamma) \), packet retransmission time \( (T) \), application latency \( (L) \) and packet loss \( (P_L) \) requirements. All these parameters are available or can be estimated in wireless communication systems and applications.

One the other hand, there is a lower bound for transmission rate for most applications. For example, live video requires a minimum throughput for content delivery \( (TP) \). Any network data rate below this \( TP \) will result in service interruption. Network data rate mainly depends on two factors. One is physical link quality and the second one is MAC layer availability of the medium. Based on these two factors, we can derive the minimum transmission rate as

\[ R \geq R_{MIN} = TP \frac{1}{1-U} \tag{8} \]

where \( R_{MIN} \) is the lower bound for transmission rate and \( U \) is the MAC layer channel utilization rate.

**Outage Capacity** [17] is the highest achievable data rate with outage probability being less than a threshold. Based on this concept, the non-utilized Outage Capacity \( C_{NUOC} \), is derived as \[8\]

\[ C_{NUOC} = B(1-U)\exp(-\frac{\gamma_{min}}{\gamma})\log_2(1+\gamma_{min}) \tag{9} \]

Using equation (5), non-utilized Outage Capacity will be simplified to

\[ C_{NUOC} = R(1-U)\exp\left(1 - \frac{2R/B}{\gamma}\right) \tag{10} \]

This equation is shown in figure 1 using \( U = 0.25, \gamma = 30db \) and \( B = 20MHz \). This figure shows that the maximum network capacity of 95Mbps can be achieved for this scenario when data rate is at 150Mbps. If transmitter uses lower or higher rates, the network performance will deteriorate. We call it optimum rate, \( R_{opt} \). Equation (10) does not consider QoS requirement for each application. Therefore, we introduce ACMAC which is designed to consider QoS requirement while optimizing NUOC. By combining the upper and lower bounds in (7) and (8) to non-utilized Outage Capacity definition in (10), we arrive at

\[ ACMAC = \max_{R} R(1-U)\exp\left(1 - \frac{2R/B}{\gamma}\right) \tag{11} \]

subject to \( R_{MIN} \leq R \leq R_{MAX} \)

Equation (11) shows a multi-purpose cross-layer formulation which can be used to optimize network performance. This formulation considers Phy layer channel quality and capacity, MAC layer channel utilization and constraints imposed by QoS related to application layer requirements. Figure 2 shows four possible scenarios for transmit rate constraint impact on the network performance. First scenario is a case where \( R_{MAX} < R_{MIN} \). This means that there is no way to satisfy QoS requirements under current network conditions. One possibility for this case to happen is when network utilization is very high resulting in high \( R_{MIN} \) value. Second scenario happens when \( R_{MIN} \leq R_{opt} \leq R_{MAX} \). In this case, rate constraint created by QoS requirements do not impact network performance because network can operate at its maximum capacity and still satisfy all QoS requirements. Third scenario happens when \( R_{MIN} \leq R_{MAX} \leq R_{opt} \). In this case, transmitter has to reduce its rate from \( R_{opt} \) to \( R_{MAX} \) to satisfy QoS requirement. This change in transmit rate results in lower network capacity but more reliable data transmission. Fourth scenario happens when \( R_{opt} \leq R_{MIN} \leq R_{MAX} \). In this case, transmitter has to increase its rate from \( R_{opt} \) to \( R_{MIN} \) to avoid congestion and be able to keep up with its ingress traffic. This condition does not necessarily result in high quality user experience but it avoids complete service disconnect in a case like live high data rate UDP video transmission where there is no congestion control at transport layer.

V. SIMULATION RESULTS

The theoretical derivation and analysis presented in previous sections have significant practical relevance. The ACMAC
criterion in equation (11), can be used for several network optimization purposes such as rate, channel and route optimization. In this section, we show how AMAC applies to rate optimization problem in wireless networks.

A. Simulation Environment

For our simulations, we use a testbed with QHS710 IEEE802.11n solution [21] with bandwidth set as 40MHz. We use four devices in our testbed, two configured as Access Point (AP) and the other two as client. One pair is used for testing and measurement and other pair is used for generating interference in the same channel. We use bandwidth of 40MHz ($BW = 40MHz$) for all devices. Ixia software [22], IxChariot, is used to generate traffic from AP to client. Ixchariot is utilized to generate three types of traffic of IPTV, VoIP and Internet data traffic. For each traffic type, we monitor same three QoS parameters (throughput, latency and packet loss) as used in section IV. These requirements are normally defined by application providers or network management entity and can vary from one case to another. In our simulations, for Internet based data traffic, best effort throughput is used where there is no specific minimum rate requirement. For VoIP and IPTV traffic, we consider high-end fixed throughput requirement of 200kbps and 15Mbps respectively. In networking, latency and packet loss requirements are usually defined for end-to-end transmission. Since in this paper, we only consider the wireless part of the network, we define wireless specific latency and PER requirements. We consider wireless latency and PER requirement for voice, video and data to be 20msec, 100msec and 500msec and $10^{-2}$, $10^{-6}$ and $10^{-7}$ respectively. Table I shows our considered requirements.

<table>
<thead>
<tr>
<th>Application</th>
<th>Throughput</th>
<th>PER</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice</td>
<td>200kbps</td>
<td>$10^{-2}$</td>
<td>20msec</td>
</tr>
<tr>
<td>Video</td>
<td>15Mbps</td>
<td>$10^{-6}$</td>
<td>100msec</td>
</tr>
<tr>
<td>Data</td>
<td>Best Effort</td>
<td>$10^{-7}$</td>
<td>500msec</td>
</tr>
</tbody>
</table>

TABLE 1

QoS REQUIREMENT USED FOR DIFFERENT APPLICATIONS

B. Simulation Results

Figure 3 shows the global optimal transmit phy rate ($R_{opt}$) and lower and upper bound rates ($R_{MIN}$ and $R_{MAX}$) at different SINR levels for applications presented in table I ($T = 10msec$). As previously mentioned, $R_{MIN}$ for data is zero. Further, $R_{MIN}$ is close to zero (less than 1 Mbps) for voice because of its low throughput requirement. However, voice has lower $R_{MAX}$ constraint than other applications and $R_{opt}$ is above this constraint (similar to third scenario in figure 2). This means that transmit rate for voice has to be reduced from the rate which achieves optimal capacity to satisfy its strict latency requirement. For video, at lower SINR levels, $R_{MAX}$ constraint is also lower than $R_{opt}$ which means lower than capacity optimal rate is required to satisfy QoS requirements. For video application at higher SINR and data application at any SINR, $R_{opt}$ falls between $R_{MIN}$ and $R_{MAX}$ similar to second scenario in figure 2, therefore system can operate at the rate which achieves maximum capacity.

![Fig. 3. $R_{opt}$ and $R_{MAX}$ for three applications with different QoS requirements](image)

![Fig. 4. ACMAC for three applications with different QoS requirements](image)

![Fig. 5. Comparison of different capacity definitions for voice application](image)

Figure 4 shows calculated ACMAC based on equation (11) for all three applications ($U = 0.5$). As it is shown, lower capacity can be achieved with voice application because of its QoS constraint. For video at lower SINR, slightly lower ACMAC is achieved. At high SINR for video and data, since there is not QoS limiting constraint, ACMAC = NUOC can be achieved. It can be concluded that rate selection methods which are not considering application layer, may achieve correct rate if QoS constraint are not strict. However, for sensitive interactive applications like voice, gaming and teleconferencing, application layer consideration is required for good Quality-of-Service.
Figure 5 shows the difference between outage capacity (Phy layer only definition) [17], NUOC (Phy and MAC layer definition) [5] and ACMAC (Phy, MAC and Application layer definition). It also shows the predicted voice throughput which can be achieved by reducing overheads from ACMAC and satisfying its QoS constraints. We have also shown the rate we get from our testbed at different locations mapped to SINR axis. We have transmitted controlled traffic from our testbed at different locations mapped to SINR axis. We have also transmitted voice at different fixed rates and found the maximum fixed rate which can achieve PER of lower than $10^{-2}$ with only two retries allowed because $L/T = 2$. These results show that ACMAC is much more accurate in predicting system performance than previous capacity definitions.

It is important to note that lack of considering higher layers in outage capacity definition may result in a significant overestimation of network capability. This overestimation can happen more frequently in dense wireless environments like office buildings or Multi-Dwelling Units (MDU) or for applications with sensitive QoS requirement like voice.

We also measured packet loss during our experiments with voice packets. We ran two test, one with rate selection mechanism based on NUOC metric and another one with rate selection based on ACMAC. Figure 6 shows the packet loss comparison for voice packets with these two different methods. As it was shown in previous figure, ACMAC based rate selection transmits data with lower rate than NUOC based rate selection. As a result, ACMAC achieves better PER performance which is within QoS requirement range (less than 1%) while NUOC significantly violates this requirement. This is showing cross layer awareness of ACMAC metric while optimizing parameters in different networking layers.

VI. CONCLUSION

A cross-layer metric considering physical, MAC and application layers called ACMAC is introduced in this paper. This metric improves accuracy of outage capacity definition by adding medium utilization awareness. It also considers application PER, latency and throughput to derive lower and upper bound constraints on network operating transmission range. Similar to capacity, ACMAC is a multi-purpose evaluation metric which can be used for cross layer optimization of different networking parameters such as transmit rate, operating channel and route optimization.

REFERENCES