Resolving 802.11 Performance Anomalies through QoS Differentiation

Hyogon Kim, *Member, IEEE,* Sangki Yun, *Student Member, IEEE,* Inhye Kang, *Member, IEEE,* and Saewoong Bahk, *Member, IEEE*

Abstract— It was recently shown that 802.11b MAC has an "anomaly" that the throughput of high bit-rate terminals in good channel condition is down-equalized to that of the lowest bit-rate peer in the network. In this letter, we analytically prove that the phenomenon can be cleanly resolved through configuring the initial contention window size inversely proportional to the bit-rate. Although contention window size variation has been used to induce throughput differentiation among equal bit-rate terminals, our work is the first to apply the technique to *different* bit-rate terminals.

Index Terms— Wireless LAN, 802.11 performance anomaly, service differentiation.

I. INTRODUCTION

IEEE 802.11 [1] supports dynamic rate shifting (DRS), which adaptively sets the nominal bit-rate of a terminal according to the channel condition. Although DRS is a measure to ensure transmission reliability under different channel conditions, it is our intuition that different bit-rates should lead to corresponding throughput differences. But a recent work by Heusse *et al.* [2] shows otherwise. Specifically, the throughput of a high-bit terminal is down-equalized to that of the lowest bit-rate peer. The first implication of this "anomaly" is the infeasibility of service differentiation, and the second, degradation of the aggregate throughput [2]. In this letter, we address the first aspect of the problem, *i.e.*, down-equalization, using service differentiation approach. But we can show that the second aspect of the anomaly, aggregate throughput drop, is also significantly alleviated through it [3].

There have been considerable efforts to create service differentiation in 802.11 wireless LAN. Aad and Castelluc-cia [4] considers three 802.11 parameters to create differentiation effects among equal bit-rate terminals. The three parameters are the initial contention window size (CW_{min}), DCF interframe space (DIFS), and the maximum frame size. The work shows the usefulness of the parameters to create differences among terminals, but falls short of showing how to create a specific target throughput or delay ratio. Zhao *et al.* go on to show CW_{min} can be utilized to create target throughput ratios among equal bit-rate terminals, by scaling

CW_{min} in proportion to the given ratio [5]. In this letter, we analytically prove that controlled throughput ratio can be precisely achieved among different bit-rate terminals using the CW_{min} control, thereby resolving the first aspect of the 802.11 performance anomaly. To the best of our knowledge, this is the first work that shows that the 802.11 performance anomaly can be resolved through CW_{min} differentiation.

II. ANALYSIS

The idea of our differentiation scheme is to set CW_{min} for each bit-rate inversely proportional to the bit-rate. (Let l denote a bit-rate in the hierarchy, which we call *class*. In this letter, we will use the example of 802.11b, so there are 4 classes: 1 Mb/s, 2 Mb/s, 5.5 Mb/s, and 11 Mb/s. First, the highest bit-rate terminals in the bit-rate hierarchy (*i.e.*, 11 Mb/s) retain the current default value, *i.e.*, $CW_{min}^{(11)} = 32$ (slots). Then CW_{min} of other bit-rate terminals is set to $CW_{min}^{(x)} = CW_{min}^{(11)} \cdot \frac{11}{x}$, where $x = 1$, 2, or 5.5. We will prove that through this simple configuration we can precisely create the throughput ratio corresponding to the bit-rate ratio. As a matter of fact, the readers should now notice, *any* throughput ratio can be created this way. For the ensuing analysis we make some simplifying assumptions as follows.

- The number of terminals in the system, N, is large.
- The maximum contention window size, CW_{max}, can be set high.
- All nodes are within the receive range of each other. This assumption excludes the hidden terminal problem.

Although these assumptions significantly simplify our analysis, they are not essential. In fact, the simulation results that corroborate the analysis are obtained in small N network. The second assumption is needed to ensure correct throughput ratio upon high contention, or equivalently, large number of backoffs [3].

Since the payload size is independent of the nominal bit-rate, the throughput ratio between different class terminals becomes the inverse ratio of the transmission times $T^{(l)}$. (The transmission time is what it takes until a class-l node to successfully transmit a packet.) Namely,

$$R_{ij} = \frac{\Theta^{(i)}}{\Theta^{(j)}} = \frac{K/T^{(i)}}{K/T^{(j)}} = \frac{T^{(j)}}{T^{(i)}} \qquad (1)$$

where R_{ij} is the throughput ratio and K is the packet size in bits. In this letter, we assume that terminals are sufficiently backlogged with data, so the packets are always transmitted
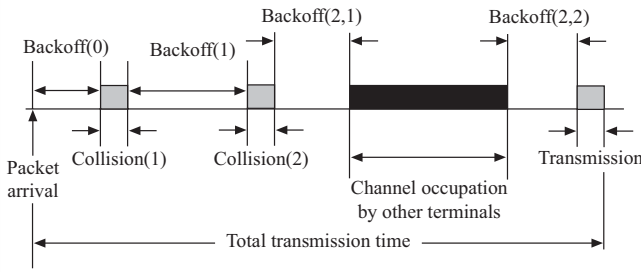
Fig. 1.   An example scenario for the model of transmission time.

in full size. To determine R_{ij}, we just need to compute $T^{(l)}$ for each class l. $T^{(l)}$ is composed of several components:

1) $E[T_s^{(l)}]$: Packet transmission time of a class l terminal. With maximal packet size assumption, $E[T_s^{(l)}] = T_s^{(l)} = K/l$.

2) $E[T_c^{(l)}]$: Average time wasted on a mangled packet transmission due to collision. It is a packet time, and does not contain the backoff times due to the collision.

3) $E[T_i^{(l)}]$: Average idle time between transmission attempts for a class l terminal. Since we assume that terminals are backlogged, this idle time is due to the backoffs for collision avoidance.

4) $E[T_o^{(l)}]$: Time of channel occupation by the terminals in classes other than l. In 802.11, when the terminal is detected busy, the backoff timer for the terminal in backoff is suspended. Therefore, this time is added to the total cost for the other terminals. Since "channel busy" could mean either successful transmission or collision, the possible collision time of other terminals is also accounted for.

Fig. 1 exemplifies our model for $T^{(l)}$. In the figure, the transmission by this terminal suffers from 2 collisions, and the third backoff time is split into two by the transmissions (and possible collisions) of other terminals. Collectively, "Backoff(0)", "Backoff(1)", "Backoff(2,1)", "Backoff(2,2)" comprise $T_i^{(l)}$. The "Transmission" duration is $T_s^{(l)}$, and "Collision(1)" and "Collision(2)" add up to $T_c^{(l)}$. Finally, the "Channel occupation by other terminals" corresponds to $T_o^{(l)}$.

With M classes, Fig. 1 tells us that

$$
\begin{aligned}
T^{(l)} & = c^{(l)} \cdot E[T_c^{(l)}] + T_s^{(l)} \\
& \sum_{k \neq l}^{M} \frac{N^{(l)}}{N^{(k)}} \cdot \frac{P_s^{(k)}}{P_s^{(l)}} (T_s^{(k)} + c^{(k)} \cdot E[T_c^{(k)}]) \\
& + (c^{(l)} + 1) \cdot E[T_i^{(l)}]
\end{aligned} \tag{2}
$$

where $c^{(l)}$ is the expected number of collisions for a class-l terminal. Here, $P_s^{(l)}$ is the probability that a class-l terminal succeeds in transmission, and is given as:

$$
P_s^{(l)} = \binom{N^{(l)}}{1} \tau^{(l)} (1 - \tau^{(l)})^{N^{(l)}-1} \prod_{k \neq l} (1 - \tau^{(k)})^{N^{(k)}} \tag{3}
$$

Thus $\frac{P_s^{(k)}}{P_s^{(l)}}$ in (2) becomes the relative number of class-k transmissions for 1 class-l transmission. The ratio $\frac{N^{(l)}}{N^{(k)}}$ in (2) is to take account of the different class population in

computing the individual total transmission time $T^{(l)}$. Note $c^{(l)}$ is the average value of a geometric random variable. Thus $c^{(l)} = \frac{1}{1-p^{(l)}} - 1 = \frac{p^{(l)}}{1-p^{(l)}}$ where $p^{(l)}$ is the class-l conditional collision probability in a slot given as:

$$
p^{(l)} = 1 - (1 - \tau^{(l)})^{N^{(l)}-1} \prod_{k \neq l} (1 - \tau^{(k)})^{N^{(k)}} \tag{4}
$$

where $N^{(l)}$ is the number of terminals in class l and $\tau^{(l)}$ is the per-slot transmission probability of class l as in Bianchi [6],

$$
\tau^{(l)} = \frac{2}{1 + CW_{min}^{(l)} + p^{(l)} CW_{min}^{(l)} \sum_{i=0}^{m} (2p^{(l)})^i} \tag{5}
$$

We can numerically compute $E[T_s^{(l)}]$, $E[T_c^{(l)}]$, and $E[T_i^{(l)}]$ in (2). In [3] we show that the throughput obtained from this numerical analysis closely matches simulation.

With a realistic model in (2), we now analytically prove that

$$
R_{ij} = \frac{\Theta^{(i)}}{\Theta^{(j)}} = \frac{T^{(j)}}{T^{(i)}} = \frac{i}{j} \tag{6}
$$

Substituting (2) for T's in (1) and considering two arbitrary classes for pairwise class-to-class throughput ratios we obtain:

$$
R_{ij} = \frac{C^{(j)} + \frac{N^{(j)}}{N^{(i)}} \cdot \frac{P_s^{(i)}}{P_s^{(j)}} C^{(i)} + \sum_{k \neq i, k \neq j} \frac{N^{(j)}}{N^{(k)}} \cdot \frac{P_s^{(k)}}{P_s^{(j)}} \cdot C^{(k)} + D^{(j)}}{C^{(i)} + \frac{N^{(i)}}{N^{(j)}} \cdot \frac{P_s^{(j)}}{P_s^{(i)}} C^{(j)} + \sum_{k \neq i, k \neq j} \frac{N^{(i)}}{N^{(k)}} \cdot \frac{P_s^{(k)}}{P_s^{(i)}} \cdot C^{(k)} + D^{(i)}} \tag{7}
$$

where $C^{(l)} = c^{(l)} \cdot E[T_c^{(l)}] + T_s^{(l)}$, and $D^{(l)} = (c^{(l)} + 1) \cdot E[T_i^{(l)}]$. Now, in (7) we have

$$
\frac{N^{(i)}}{N^{(j)}} \cdot \frac{P_s^{(j)}}{P_s^{(i)}} = \frac{\tau^{(j)} \cdot (1 - \tau^{(i)})}{\tau^{(i)} \cdot (1 - \tau^{(j)})} \tag{8}
$$

But if we assume CW_{max} can be set high, m becomes sufficiently large in (5). Then,

$$
\begin{aligned}
\frac{\tau^{(j)} \cdot (1 - \tau^{(i)})}{\tau^{(i)} \cdot (1 - \tau^{(j)})} & = \frac{CW_{min}^{(i)} \cdot \frac{1-p^{(i)}}{1-2p^{(i)}} - 1}{CW_{min}^{(j)} \cdot \frac{1-p^{(j)}}{1-2p^{(j)}} - 1} \\
& \approx \frac{CW_{min}^{(i)}}{CW_{min}^{(j)}} = \frac{j}{i} \tag{9}
\end{aligned}
$$

since $1 - p > 1 - 2p$ and $CW_{min}^{(l)} >> 1$ (the smallest $CW_{min}^{(l)}$ is 31 (slots) for $l = 11$(Mb/s)). Likewise, for $l = i$ or j, $\sum_{k \neq i, k \neq j} \frac{N^{(l)}}{N^{(k)}} \cdot \frac{P_s^{(k)}}{P_s^{(l)}} \cdot C^{(k)} = \frac{k}{l}$. Reflecting this, (8), and (9) to (7), we obtain

$$
R_{ij} = \frac{i}{j} \cdot \frac{j \cdot C^{(j)} + i \cdot C^{(i)} + \sum_{k \neq i, k \neq j} k \cdot C^{(k)} + j \cdot D^{(j)}}{i \cdot C^{(i)} + j \cdot C^{(j)} + \sum_{k \neq i, k \neq j} k \cdot C^{(k)} + i \cdot D^{(i)}} \tag{10}
$$

Meanwhile,

$$
\begin{aligned}
\frac{D^{(j)}}{D^{(i)}} & = \frac{(c^{(j)} + 1) \cdot E[T_i^{(j)}]}{(c^{(i)} + 1) \cdot E[T_i^{(i)}]} \\
& \approx \frac{E[T_i^{(j)}]}{E[T_i^{(i)}]} = \frac{\sigma \cdot \frac{1-\tau^{(j)}}{\tau^{(j)}}}{\sigma \cdot \frac{1-\tau^{(i)}}{\tau^{(i)}}} = \frac{i}{j}
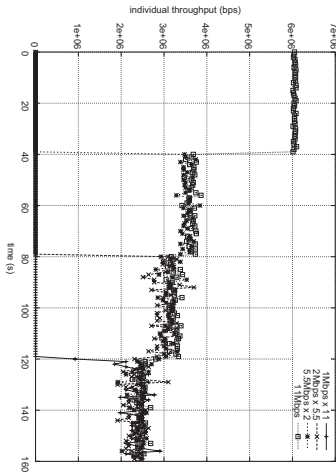\end{aligned}
$$

Fig. 2.   Scaled throughput of each class

by recycling the derivation in (9) and by $c^{(i)} \approx c^{(j)}$. (Here, σ is the length of a slot time.) Then (10) becomes

$$R_{ij} = \frac{i}{j} \cdot \frac{j \cdot C^{(j)} + i \cdot C^{(i)} + \sum_{k \neq i, k \neq j} k \cdot C^{(k)} + j \cdot D^{(j)}}{i \cdot C^{(i)} + j \cdot C^{(j)} + \sum_{k \neq i, k \neq j} k \cdot C^{(k)} + j \cdot D^{(j)}}$$

$$= \frac{i}{j} \qquad (11)$$

Therefore, the throughput ratio is directly proportional to the nominal bit-rate ratio under inversely proportional CW_{min} configuration.

Now we validate (11) through simulation in ns-2 [7]. Fig. 2 shows a scenario where 11 Mb/s, 5.5 Mb/s, 2 Mb/s, and then 1 Mb/s terminals are added one by one, 40 seconds apart. For the bit-rates other than 11 Mb/s, we scaled their throughput in the

graph for easier comparison. The figure shows that the ratio is 11:5.5:2:1 as we intended (i.e. the scaled ratio is 1:1:1:1). Therefore, we demonstrated that the differentiation by CW_{min} control achieves the target throughput distribution.

In [3], we also demonstrate that the service differentiation improves the system throughput as well, addressing the second aspect of the 802.11 anomaly.

III. CONCLUSION

CW_{min} is expected to be a configurable parameter in 802.11e, the extension for QoS provision in 802.11 wireless networks. In this letter, we analytically proved that controlling CW_{min} can resolve the "802 performance anomaly." But more generally, this method works for any throughput ratio between wireless terminals for QoS differentiation. Issues with other differentiation mechanisms, channel access delay, and TCP are also discussed in an unabridged version [3].

REFERENCES

[1] ANSI/IEEE, "802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2000.
[2] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance Anomaly of 802.11b," in proceedings of IEEE Infocom, 2003.
[3] H. Kim, S. Yoon, and I. Kang, "Resolving 802.11 performance anomaly through service differentiation," Technical Report, KU-CS-TR-04-0382, Korea University, Nov. 2004. Also available at http://ubiquitous.korea.ac.kr/n-ano.pdf.
[4] I. Aad and C. Castelluccia, "Differentiation mechanism for IEEE 802.11," in proceedings of IEEE Infocom, 2001.
[5] J. Zhao, Q. Zhang, Z. Guo, W. Zhu, and Y.-Q. Zhang, "Throughput and QoS optimization in IEEE 802.11 WLAN," 3Gwireless2002, 2002.
[6] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Area in Communications*, March 2000, 18(3): 535-547.
[7] The ns-2 simulator. Available at http://www.isi.edu/nsnam/ns/.