Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

# A novel MAC mechanism to resolve 802.11 performance anomaly*

CHENG Yan-hong†, LI Zhi-shu†‡, XING Jian-chuan, ZHU Li
(*School of Computer Science, Sichuan University, Chengdu 610065, China*)
†E-mail: bjcyh@263.net; Lzshu5@yahoo.com.cn

Received Mar. 26, 2007; revision accepted June 11, 2007

**Abstract:** In the 802.11b networks, the guarantee of an equal long-run channel access probability causes performance anomaly in a multi-rate wireless cell. Much interest has been involved in this issue and many effective mechanisms have been proposed. The usual MAC layer solutions include the initial contention window adaptation, the maximum transfer unit size adaptation and the packet bursting. In this paper, we propose a novel approach which introduces a new parameter called the transmission probability $p_t$ to the legacy protocol. By adjusting $p_t$ according to the transmission rate, the proposed scheme can solve the performance anomaly problem cleanly. Throughput analysis and performance evaluation show that our scheme achieves significant improvement in the aggregate throughput and the fairness.

**Key words:** IEEE 802.11, WLAN, Transmission probability, Performance anomaly
**doi:**10.1631/jzus.2007.A1573 **Document code:** A **CLC number:** TN92

## INTRODUCTION

In recent years, the IEEE 802.11 WLAN standard has become increasingly popular as one of the main wireless communication technologies, which in turn has fostered the research in improving its performance. In accessing a shared wireless channel, fairness is an important issue, which becomes even more pronounced for multi-rate WLANs.

There are some reasons for the diversity of data rate in a wireless cell. First, 802.11b, the most widely used version of 802.11, supports the technology of dynamic rate shifting (DRS), which provides nodes with the capability of adapting the transmission rate according to the channel condition. For example, when channel condition is good, the transmission rate can be set to the maximum value of 11 Mbps; when channel condition becomes worse, e.g., in the scenarios where the interference occurs or the communication peers are out of the receiving range of each other, the transmission rates might be tuned to a lower level (e.g., 5.5, 2 or 1 Mbps). Therefore, 802.11b may diversify the data rate of stations sharing the same wireless channel. Secondly, the use of compatible technologies of different 802.11 versions in a wireless cell could also cause stations to transmit at different data rates. For example, the maximum data rate is 11 Mbps for 802.11b but 52 Mbps for 802.11a and 802.11g.

However, Heusse *et al.*(2003) pointed out that the diversity of data rate in a wireless cell could lead to the performance anomaly in IEEE 802.11 networks. Namely, the throughput of high bit-rate nodes reduces to that of the lowest bit-rate peer. The root cause of this anomaly lies in the basic CSMA/CA channel access protocol which provides each active node with fair channel access. In the 802.11b protocol, each node experiences approximately the same number of transmissions in spite of the amount of time required to transmit a packet. Thus low bit-rate node will occupy the wireless channel much longer than high bit-rate node. Consider a scenario where two nodes in a wireless cell send equal-size packets to the access point (AP) at 1 Mbps and 11 Mbps, respectively. If the protocol overhead is ignored, the channel time spent by the slow node will be 11 times as much as that

---

‡ Corresponding author

* Project supported by the National Mid- and Small-Scale Enterprise Technical Innovation Fund of China (No. 06C26225101730) and the Sichuan Provincial Key Science and Technology Program, China (No. 05GG021-003-2)

spent by the fast node. As a consequence, these two nodes achieve the same throughput even though their data rates are greatly different and the aggregate throughput suffers a significant degradation.

The performance anomaly indicates that the throughput-based fairness widely accepted in single-rate WLANs cannot suit multi-rate WLANs well. As pointed out in (Tan and Guttag, 2004; Li and Soung, 2005), the above inefficiencies can be addressed by adopting the notion of the time-based fairness, which yields a good trade-off between fairness and efficiency in a multi-rate WLAN. Based on this idea, a novel mechanism called TPA (transmission probability adaptation) is proposed in this paper with the introduction of an additional parameter—the transmission probability $p_t$ to the 802.11b protocol. By configuring $p_t$ according to the data rate, the mechanism provides each node with an equal share of the wireless channel occupation time to eliminate the performance anomaly.

The remainder of this paper is organized as follows. Section 2 reviews the existing works which resolve the performance anomaly through tiny modification to IEEE 802.11. Section 3 introduces the novel resolution to performance anomaly called TPA and validates its effectiveness through a detailed theoretical analysis. In Section 4, we present simulation results by means of the ns-2 simulator and compare TPA with some other schemes. Section 5 concludes this paper. Furthermore, the access probability for TPA is derived in the Appendix.

## RELATED WORK

There are a significant number of research efforts focusing on the issue of fairness in multi-rate WLANs. Many solutions performed at different protocol layers have already been proposed in the literature, among which some adopt multi-queue scheduling algorithms based on the leaky bucket scheme, e.g. (Tan and Guttag, 2004; Garroppo *et al.*, 2007), requiring no modification to the underlying MAC protocol. However, since the source of the performance anomaly is rooted in the basic CSMS/CA, it seems most natural to seek to resolve this issue at the MAC layer itself. Like the TPA scheme presented in this paper, there are three classes of schemes aiming at solving the performance anomaly problem by introducing tiny modification to the IEEE 802.11 standard, i.e., initial contention window adaptation approach, maximum transfer unit size adaptation, and packet bursting approach.

### Initial contention window adaptation approach

The first class of solution achieves the time-based fairness by adjusting the initial contention window size $CW_{\min}$. Kim *et al.*(2005) configured $CW_{\min}$ inversely proportional to the data rate and analytically proved the effectiveness of this scheme. In fact, since the transmission overhead has not been considered by this scheme, the performance anomaly cannot be cleanly resolved. A similar approach is proposed by Iannone *et al.*(2005), but obtaining the time-based fairness is not its only goal. In order to further improve channel utilization, this approach adjusts the $CW_{\min}$ so that the number of idle slots perceived by a station can converge to the target value. The station with lower data rate then scale its $CW_{\min}$ according to the ratio of the maximum available rate to its data rate. As the former proposal does, this scheme also ignores the protocol overhead.

### Maximum transfer unit size adaptation

Adjusting the maximum transfer unit (MTU) size is the simplest approach. Yoo *et al.*(2005) reduced the performance anomaly by assigning the MTU size proportionally depending on the data rate. Simulation study shows this scheme achieves an improvement in the throughput and the fairness. Dunn *et al.*(2006) also proposed a similar approach by introducing an MTU discovery process to determine the packet size according to the data rate. Nevertheless, the proposed solution is efficient only for the stations with the highest data rate. For slow stations, due to the overhead and backoffs introduced by the fragmentation, the achieved throughput is below the ideal value.

### Packet bursting approach

As defined in 802.11e, packet bursting is used to achieve equal transmission opportunity (TXOP) for all the stations which can lead to equal channel occupancy time directly. Sadeghi *et al.*(2002) proposed a solution called OAR (Opportunistic Auto Rate) protocol, which allows the station with high data rate to transmit multiple packets in order to take advantage of the favorable channel condition. The number of successive packets is allocated to satisfy the condition that each station can obtain the same TXOP lengths. PAS (Performance Anomaly Solution) is another solution proposed by Razafindralambo and Lannone (2006), which also belongs to packet bursting approach but in a dynamic way. Each station should sense the channel continuously to discovery the channel occupancy time, which exceeds the transmission time of its packets; then it can aggregate packets to reach the same channel occupancy time. PAS improves the short time unfairness introduced by successive transmission of the same node. The packet bursting approach can improve the throughput and fairness index in one hop case because of the reduction of the transmission overhead. However, the solution could suffer serious fairness degradation in multi-hop case (Yoo *et al.*, 2005).

Different from these conventional technologies, we propose a novel mechanism called TPA, with the expectation of contributing a new idea to overcome the performance anomaly problem. This scheme introduces the transmission probability $p_t$ to 802.11 DCF, thus needs a small variation of 802.11 DCF. Analysis and simulation results show this scheme can improve throughput and fairness significantly. Our scheme can be put into effect flexibly since $p_t$ is independent of the 802.11 system and can be set to an arbitrary value equal to or less than 1 as needed.

## RESOLVING 802.11 PERFORMANCE ANOMALY

The proposed mechanism adopts the idea of *p*-persistent CSMA which differs from 802.11 in that the backoff algorithms operates, and still maintains the binary exponential backoff (BEB) algorithm adopted by the standard protocol. The modification to IEEE 802.11 in our scheme is described as follows: after a DCF interframe space (DIFS) as backoff timer reaches zero, the station initiates a transmission at the probability of $p_t$, while keeping the previous contention window size (*CW*) unchanged, and starts the next backoff process at the probability of $1-p_t$. The only difference between our scheme and IEEE 802.11 is the value of $p_t$ after gaining the channel access opportunities: the former remains 1 all the time and the latter can be set to any value less than or equal to 1 as needed. The probability that a station transmits a packet in a randomly chosen slot time is called access probability in this paper. Let $\tau$ and $\tau'$ respectively represent the access probability for our scheme and for the IEEE 802.11 under the saturation conditions, the relation between $\tau$ and $\tau'$ (see the Appendix) is $\tau=p_t\tau'$. The simple proportional relation indicates that the access probability needed can be readily achieved by the adaptation of $p_t$. In this section, we demonstrate that the performance anomaly can be solved by tuning $p_t$ according to the data rate.

As described in (Kim *et al.*, 2005), the performance anomaly includes two implications: the infeasibility of service differentiation and the degradation of the aggregate throughput. Therefore, a complete solution should not only provide service differentiation but also improve the aggregate throughput. In this section, we argue the achievements of our scheme in both aspects through throughput analysis. We assume a general wireless network case in which *N* nodes of different bit rates compete for the wireless channel. According to the data rate, the nodes are divided into *M* classes, each of which has $N(i)$ nodes and is associated with $p_t(i)$. We use some symbol conventions in (Kim *et al.*, 2005) and make some assumptions to simplify our analysis:

(1) All nodes operate in the ideal channel (i.e., no hidden terminal or capture) and saturation conditions (i.e., the transmission queue of each node is always nonempty).

(2) The number of stations in the wireless cell is large and the average packet payload sizes of all nodes are the same.

(3) The maximum contention window size ($CW_{\max}$) can be set high enough to ensure the correct throughput ratio at high loads.

### Transmission cycle of a given node

For the convenience of description, we first investigate the transmission cycle of a given node called $T(k)$. $T(k)$ is the overall time required by a given class-*k* terminal to successfully transmit a packet. $T(k)$ consists of four components:

$$T(k)=T_s(k)+C(k)T_c(k)+[C(k)+1)]T_i(k)+T_o(k), \quad (1)$$

where $T_s(k)$, $T_c(k)$, $T_i(k)$ and $T_o(k)$ are respectively described as follows:

$T_s(k)$ is the time for one class-*k* successful transmission, which incorporates not only the delivering time of the data packets, but also the overhead such as protocol header, ACK, IFS, propagation delays, etc.

$T_c(k)$ is the time wasted on the channel because of one class-$k$ collision. According to 802.11, the average collision number $C(k)$ of a given class-$k$ node during $T(k)$ can be computed as a function of conditional collision probability $p(k)$:

$$C(k) = p(k)/[1 - p(k)], \tag{2}$$

where $p(k)$ is the class-$k$ conditional collision probability in a slot given as:

$$p(k) = 1 - [1 - \tau(k)]^{N(k)-1} \prod_{l \neq k; l=1}^{M} [1 - \tau(l)]^{N(l)}, \tag{3}$$

$\tau(k)$ is the class-$k$ access probability in a slot, and can be expressed as (see the Appendix):

$$\tau(k) = 2p_t(k) \bigg/ \left\{ 1 + W + p(k)W \sum_{i=0}^{m-1} [2p(k)]^i \right\}. \tag{4}$$

$T_i(k)$ is the idle time between two transmissions of a given class-$k$ node. Since the transmission number of a given class-$k$ node in $T(k)$ is $C(k)+1$, the total idle time spent by the node is $[C(k)+1]T_i(k)$. Similarly, $T_i(k)$ is the average value of a geometric random variable:

$$T_i(k) = \sigma[1 - \tau(k)]/\tau(k), \tag{5}$$

where $\sigma$ is the slot time size.

$T_o(k)$ is the time spent by all nodes except for the given node to transmit packets. Since the same class nodes have the same channel access opportunity, each class-$k$ node also experiences one successful transmission in the time $T(k)$, then $T_o(k)$ can be written as:

$$\begin{aligned} T_o(k) = {} & [N(k) - 1][T_s(k) + C(k)T_c(k)] + \\ & \sum_{l \neq k; l=1}^{M} \frac{P_s(l)}{P_s(k)} N(l)[T_s(l) + C(l)T_c(l))], \end{aligned} \tag{6}$$

where the probability $P_s(k)$ that a given class-$k$ node succeeds in transmission is given by the probability that this node transmits a packet over the channel exactly:

$$P_s(k) = \tau(k)[1 - \tau(k)]^{N(k)-1} \prod_{l \neq k; l=1}^{M} [1 - \tau(l)]^{N(l)}. \tag{7}$$

Finally, from the above equations, we obtain the expression of the transmission cycle of a given node:

$$\begin{aligned} T(k) = {} & N(k)[T_s(k) + C(k)T_c(k)] + [C(k) + 1]T_i(k) + \\ & \sum_{l \neq k; l=1}^{M} \frac{P_s(l)}{P_s(k)} N(l)[T_s(l) + C(l)T_c(l)]. \end{aligned} \tag{8}$$

Based on Eq.(8), we prove that the proposed scheme can support service differentiation and improve the aggregate throughput in the next two subsections.

### Creating service differentiation

In this subsection we observe the impact of our scheme on the throughput allocated to each node with different data rates. Let $p_t(i)$ and $p_t(j)$ be the transmission probability of class $i$ and class $j$, respectively. We compute the throughput ratio of a class-$i$ traffic to a class-$j$ traffic:

$$\begin{aligned} S_{ij} & = \frac{L/T(i)}{L/T(j)} = \frac{T(j)}{T(i)} = \\ & \frac{N(j)M(j) + \sum_{k \neq j; k=1}^{M} \frac{P_s(k)}{P_s(j)} N(j)M(k) + [C(j)+1]T_i(j)}{N(i)M(i) + \sum_{k \neq i; k=1}^{M} \frac{P_s(k)}{P_s(i)} N(i)M(k) + [C(i)+1]T_i(i)}, \end{aligned} \tag{9}$$

where $M(l)=T_s(l)+C(l)T_c(l)$.

Under the assumed conditions, we have $\tau \ll 1$ and the condition collision probability of each node is equal approximately, the ratio of $P_s(k)$ to $P_s(j)$ is:

$$\frac{P_s(k)}{P_s(j)} = \frac{\tau(k)[1 - \tau(j)]}{\tau(j)[1 - \tau(k)]} \approx \frac{\tau(k)}{\tau(j)} \approx \frac{p_t(k)}{p_t(j)}. \tag{10}$$

Meanwhile, from Eq.(2) and Eq.(5), we have:

$$\frac{[C(i)+1]T_i(i)}{[C(j)+1]T_i(j)} = \frac{[1 - p(j)][1 - \tau(i)]\tau(j)}{[1 - p(i)][1 - \tau(j)]\tau(i)} \approx \frac{p_t(j)}{p_t(i)}. \tag{11}$$

By substituting Eqs.(10) and (11) into Eq.(9), the throughput ratio can be simplified as Eq.(12).

Eq.(12) shows the throughput ratio is proportional to the transmission probability ratio. This property can be utilized to support service differenttiation. For example, if some nodes need supporting real-time application or have better channel quality, we can set bigger $p_t$ for them and smaller $p_t$ for others.

$$S_{ij}=\frac{N(j)M(j)+N(i)M(i)\dfrac{p_{\mathrm{t}}(i)}{p_{\mathrm{t}}(j)}+\sum\limits_{k\neq i,j;k=1}^{M}\dfrac{p_{\mathrm{t}}(k)}{p_{\mathrm{t}}(j)}N(k)M(k)+[C(j)+1]T_{\mathrm{i}}(j)}{N(i)M(i)+N(j)M(j)\dfrac{p_{\mathrm{t}}(j)}{p_{\mathrm{t}}(i)}+\sum\limits_{k\neq i,j;k=1}^{M}\dfrac{p_{\mathrm{t}}(k)}{p_{\mathrm{t}}(i)}N(k)M(k)+\dfrac{p_{\mathrm{t}}(j)}{p_{\mathrm{t}}(i)}[C(j)+1]T_{\mathrm{i}}(j)}\approx\frac{p_{\mathrm{t}}(i)}{p_{\mathrm{t}}(j)}\ . \quad (12)$$

Then the nodes with big $p_{\mathrm{t}}$ can get more channel access opportunities than those with small $p_{\mathrm{t}}$. In this way, the bandwidth can be allocated on demand.

**Improving the aggregate throughput**

Since the throughput ratio is proportional to the transmission probability ratio, if we assign larger $p_{\mathrm{t}}$ to the fast nodes and smaller $p_{\mathrm{t}}$ to the slow nodes, then the fast nodes will obtain more channel occupancy time and thus achieve more throughput for our scheme than for the standard protocol, as a result the aggregate throughput is improved.

Let us observe the aggregate throughput of two systems: one uses the standard protocol, and the other uses the proposed scheme. According to Eq.(8), in the time $T(k)$, the total number of successful packets transmitted by all nodes is $\sum_{l\neq k;l=1}^{M}N(l)P_{\mathrm{s}}(l)/P_{\mathrm{s}}(k)$ $+N(k)$, then the aggregate throughput is:

$$S=\frac{N(k)\left[1+\sum\limits_{l\neq k;l=1}^{M}\dfrac{N(l)}{N(k)}\dfrac{P_{\mathrm{s}}(l)}{P_{\mathrm{s}}(k)}\right]L}{N(k)M(k)+[C(k)+1]T_{\mathrm{i}}(k)+\sum\limits_{l\neq k;l=1}^{M}\dfrac{P_{\mathrm{s}}(l)}{P_{\mathrm{s}}(k)}N(l)M(l)}. \quad (13)$$

When the transmission probabilities of all classes are the default value of 1, we obtain the throughput $S'$ of 802.11 protocol:

$$S'=\frac{\sum_{l=1}^{M}N(l)L}{N(k)M'(k)+[C'(k)+1]T_{\mathrm{i}}'(k)+\sum\limits_{l\neq k;l=1}^{M}N(k)M'(l)}. \quad (14)$$

In order to compare $S$ with $S'$, we first assume the class-$k$ nodes have the fastest bit-rate, thus $p_{\mathrm{t}}(k)=1$ and $p_{\mathrm{t}}(l)<1$ when $l\neq k$. Then, we can make the approximation that $C(l)\approx C'(l)$ under the assumed conditions. For the convenience of comparison, we use symbol $T$ to represent the average time spent by any node to transmit a packet successfully. The relation between $S$ and $T$ is: $S=L/T$. From Eqs.(13) and (14), $T$ and $T'$ can be respectively expressed as:

$$T=\left\{N(k)M(k)+[C(k)+1]T_{\mathrm{i}}(k)+\sum_{l\neq k;l=1}^{M}\frac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}N(l)M(l)\right\}\cdot\left\{N(k)+\sum_{l\neq k;l=1}^{M}N(l)\frac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}\right\}^{-1}, \quad (15)$$

$$T'=\frac{N(k)M(k)+[C(k)+1]T_{\mathrm{i}}'(k)+\sum\limits_{l\neq k;l=1}^{M}N(l)M(l)}{\sum_{l=1}^{M}N(l)}. \quad (16)$$

From Eq.(15), $T$ is the sum of three terms. Let $T_1$, $T_2$ and $T_3$ represent each term of $T$. Similarly, $T_1'$, $T_2'$ and $T_3'$ represent each term of $T'$.

$T_1$ and $T_1'$ are the transmission time spent by class-$k$ nodes. Subtracting $T_1'$ from $T_1$ and noting that $p_{\mathrm{t}}(l)<p_{\mathrm{t}}(k)$, we have:

$$T_1-T_1'=\frac{N(k)M(k)\sum\limits_{l\neq k;l=1}^{M}N(l)\left[1-\dfrac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}\right]}{\sum\limits_{l=1}^{M}N(l)\left[N(l)+\sum\limits_{l\neq k;l=1}^{M}N(l)\dfrac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}\right]}>0, \quad (17)$$

which indicates that class-$k$ nodes receive more channel utilization occupation time for the proposed scheme than for the standard one.

$T_2$ and $T_2'$ are the total idle time experienced by the given node during the transmission cycle. The small difference between $T_2$ and $T_2'$ is negligible under the consumed conditions.

$T_3$ and $T_3'$ are the transmission time spent by all nodes except for the class-$k$ peers. We observe the comparison between $T_3$ and $T_3'$:

$$T_3-T_3'=\frac{N(k)\sum\limits_{l\neq k;l=1}^{M}N(l)M(l)\left[\dfrac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}-1\right]}{\sum\limits_{l=1}^{M}N(l)\left[N(l)+\sum\limits_{l\neq k;l=1}^{M}N(l)\dfrac{p_{\mathrm{t}}(l)}{p_{\mathrm{t}}(k)}\right]}<0, \quad (18)$$

which indicates that slow bit-rate nodes receive a smaller amount of channel occupation time for the proposed scheme than for the standard one.

When the number of stations is large and the access probability is small, the collisions experienced by each station are approximately the same. Thus $C(l)\approx C(k)$ when $l\neq k$ and the following inequality is satisfied:

$$M(l)-M(k)=T_{\rm s}(l)+C(l)T_{\rm c}(l)-T_{\rm s}(k)-C(k)T_{\rm c}(k) >0, \quad l\neq k. \tag{19}$$

Therefore, combining Eqs.(17)~(19), we readily obtain:

$$T_1-T_1'+T_2-T_2'+T_3-T_3'<0\Rightarrow T<T'\Rightarrow S>S'. \tag{20}$$

Eqs.(17) and (18) clearly explain why our scheme achieves higher throughput than 802.11 and confirm that the slower the bit-rate and the smaller the transmission probability associated with the slower node, the higher throughput the proposed scheme achieves.

**Analysis of time-based fairness**

Since $p_{\rm t}$ directly affects the channel capacity allocation, the key problem is the assignment of $p_{\rm t}$ in accordance with the data rate. Let $R$ be the transmission probability ratio of a slow node to the fastest node in a wireless cell. If $R$ is too small, the slow node might be starved although the aggregate throughput could be improved significantly. On the contrary, if $R$ is too large, the performance anomaly cannot be solved cleanly. The assignment of $p_{\rm t}$ relates to two problems: firstly, since the source of the performance anomaly lies in the throughput-based fairness, a new fairness criterion should be formulated for multi-rate WLANs; Secondly, the derivation of optimal $p_{\rm t}$ corresponds to the new fairness criterion.

The notion of time-based fairness is proposed in (Tan and Guttag, 2004; Li and Soung, 2005) to improve performance in multi-rate WLANs by providing each competing node an equal share of the wireless channel occupancy time. The time-based fairness can guarantee that no node achieves worse throughput than it would do in a single-rate WLAN in which all competing nodes are running at their own rates (Tan and Guttag, 2004). We can evaluate the fairness performance of a solution according to this feature.

Based on the time-based fairness, we now derive the optimal value of $p_{\rm t}$. Eq.(8) has given the expression of $T(k)$ and the transmission cycle of a given class-$k$ node. In $T(k)$, the channel occupation time for a class-$k$ node is $M(k)$, while that for a class-$l$ node is $P_{\rm s}(l)M(l)/P_{\rm s}(k)$. We can observe that the channel occupancy time for each node is approximately equal when the following equation is satisfied:

$$p_{\rm t}(l)/p_{\rm t}(k)\approx M(k)/M(l). \tag{21}$$

Namely, by configuring $p_{\rm t}$ inversely proportional to the overall transmission time for a packet (including the retransmission attempt), a node achieves an equal amount of the shared channel resource, whatever its data rate is.

Ignoring the protocol overhead (e.g., physical and packet headers, acknowledgements, sensing the channel, and backoffs), we can simplify Eq.(21) by assigning $p_t$ proportional to the data rate. However, this simplification is not a good approximation for practical scenarios since the constant protocol overhead occupies a great proportion of the overall time for one successful transmission. Assuming the basic access mechanism in the 802.11b protocol, we now compute the value of the transmission probability ratio.

Since $M(k)=T_{\rm s}(k)+C(k)T_{\rm c}(k)$ [$T_{\rm s}(k)$ and $T_{\rm c}(k)$ were given in (Bianchi, 2000)], we give the relations in more detail:

$$T_{\rm s}(k)=t_{\rm PHY\_header}+t_{\rm MAC\_header}+t_{\rm p}+SIFS+\delta+ t_{\rm PHY\_header}+t_{\rm ACK}+DIFS+\delta, \tag{22}$$

$$T_{\rm c}(k)=t_{\rm PHY\_header}+t_{\rm MAC\_header}+t_{\rm p}+DIFS+\delta, \tag{23}$$

where $t_{\rm p}$ is the transmission time required for the MAC layer payload which includes the protocol headers attached by the protocol layers above MAC (e.g., 20-byte IP header), $\delta$ is the propagation delay. Each data frame and each ACK frame require a physical layer header, transmitted at 1 Mbps. ACK frames are transmitted at either 1 or 2 Mbps. Table 1 shows each fraction of time required to transmit a 1000-byte UDP packet, assuming a DSSS system parameter and 2-Mbps bit rate for ACK frame transmission.

From Table 1, we can observe that only packet header except for the PHY layer header and payload are transmitted at the nominal data rate (e.g., 1 Mbps or 11 Mbps). The constant overhead for a successful transmission includes the ACK frame, physical layer header, SIFS, DIFS and $\delta$, which is independent of the data rate. The value of $T_s(k)$ is 1266 μs, and the constant overhead adds up to 504 μs. The latter is about one third of the former, thus cannot be ignored.

**Table 1 Each fraction of time required to transmit a 1000-byte UDP packet for a classical DSSS physical layer**

| Parameters | Length (bits) | Data rate (Mbps) | Time (μs) |
|---|---|---|---|
| Physical layer header | 192 | 1 | 192.0 |
| Packet header | 384 | 11 | 34.9 |
| Packet payload | 8000 | 11 | 727.3 |
| Acknowledgement | 112 | 2 | 56.0 |
| Propagation delay | – | – | 2.0 |
| SIFS | – | – | 10.0 |
| DIFS | – | – | 50.0 |

Since the number of unsuccessful transmissions is far less than that of successful transmissions under the assumed conditions, we can make an approximation to simplify Eq.(21):

$$p_t(l)/p_t(k) \approx T_s(k)/T_s(l), \tag{24}$$

which indicates the time-based fairness can be achieved through configuring $p_t$ inversely proportional to the time required for a successful transmission when the basic access mechanism is adopted. In common scenarios where the offered load is not very high, $p_t$'s for the fastest nodes in a wireless cell are set to the default value 1. In fact, $p_t$ can be adjusted according to the contention level so that the optimal aggregate throughput can be achieved. The detailed analysis of the related things is outside the scope of this paper. Since the fastest node has the maximum transmission probability, $p_t$ for a node transmitted at a lower data rate can be computed from Eq.(24). For example, consider two nodes that transmit packets at data rates 11 Mbps and 1 Mbps, respectively. With the parameters in Table 1, we can easily compute that the transmission probability ratio of the fast node to the slow node is 7. However, for the simplified scheme, which assigns $p_t$ proportional to the data rate, the transmission probability ratio is 11. The latter might achieve higher throughput but poorer time-based fairness since it reduces the channel occupancy time for slow nodes.

## SIMULATION RESULTS

We conduct simulations based on ns-2 to investigate the performance of the TPA scheme. The adopted topology comprises a number of nodes in an ad hoc setting. All nodes are located within a basic service set, and they are static in the simulations. It is assumed as an ideal channel (i.e., no hidden terminals or capture) and all the nodes can hear each other directly. The number of nodes is even, in which half of the nodes are senders and the other half are receivers. One connection is set up between per pair nodes for data transmission. Thus the number of flows is half of the number of nodes. All of the tests are performed using UDP saturated traffic under the basic access mechanism. Unless otherwise specified, each packet contains 1000 bytes of data. The DSSS PHY layer parameters in Table 1 are also used in the simulations. Each experiment runs for 100 s.

The main performance metrics of interest are fairness index, throughput and collision rate. Collision rate is expressed as the ratio of the number of collisions to the number of transmissions. We adopt the Jain fairness index in (Jain, 1999), defined as:

$$f = \left( \sum_{i=1}^{n} s_i / s_i^* \right)^2 \bigg/ \left[ n \sum_{i=1}^{n} (s_i / s_i^*)^2 \right], \tag{25}$$

where $s_i$ is the throughput achieved by flow $i$, $n$ is the number of flows, $s_i^*$ is the throughput achieved by flow $i$ when all the flows in the wireless network are emitted at the same data rate, which is called the reference throughput of flow $i$ in this paper.

### Impact of the rate difference

In the simulation, we compare TPA and DCF in the case where only two flows are included in a wireless cell, one at $x$ Mbps ($x$=1, 2 or 5.5) and the other at 11 Mbps. For TPA scheme, we first compute the parameter $p_t$ from Eq.(24). The achieved throughput of each node, the aggregate throughput, the collision rate and the Jain fairness index for DCF and TPA are given in Table 2.

As shown in Table 2, when using the DCF scheme, a flow of any transmission rate experiences the same throughput so that the aggregate throughput exhibits a severe degradation. As the data rate of slow node decreases, performance anomaly becomes more

**Table 2 Comparison between DCF and TPA**

| $x$ (Mbps) | Types | $Thr_s$ (kbps) | $Thr_f$ (kbps) | Total (kbps) | Collision rate (%) | Fairness index (%) |
|---|---|---|---|---|---|---|
| 1 | REF | 426.738 | 2705.277 | 3132.015 | – | – |
| | DCF | 724.217 | 709.816 | 1434.033 | 6.2 | 0.651 |
| | TPA | 367.138 | 2881.744 | 3248.882 | 1.5 | 0.989 |
| 2 | REF | 795.505 | 2705.277 | 3500.782 | – | – |
| | DCF | 1216.051 | 1195.090 | 2411.141 | 6.1 | 0.767 |
| | TPA | 701.633 | 2849.494 | 3551.127 | 2.6 | 0.992 |
| 5.5 | REF | 1762.414 | 2705.277 | 4467.691 | – | – |
| | DCF | 2142.206 | 2109.084 | 4251.290 | 5.9 | 0.954 |
| | TPA | 1668.712 | 2762.360 | 4431.072 | 4.7 | 0.999 |

Only two flows are included in a wireless cell, one at $x$ Mbps, and the other at 11 Mbps. REF: the reference throughput. $Thr_s$ and $Thr_f$ are the throughput achieved by the slow flow and fast flow, respectively

serious since the channel occupancy time for the slow node increases.

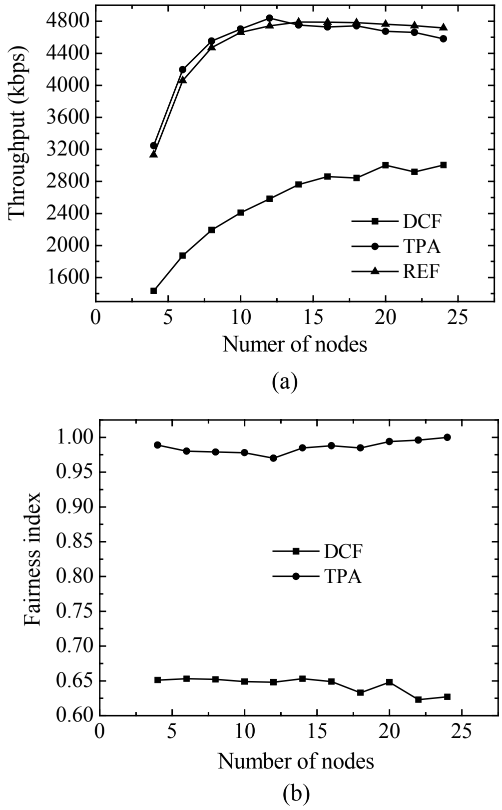
As the analysis in the previous section shows, when using the TPA scheme, the achieved throughput of each flow in each case roughly equals the corresponding reference throughput. Nevertheless, the throughput achieved by the slow flow for TPA is always slightly lower than its reference throughput, while the throughput achieved by the fast flow for TPA is always slightly higher than its reference throughput, because the slow flow experiences more collisions and the fast flow for TPA experiences fewer collisions when using TPA scheme.

The throughput of the fast node almost remains the same, independent of the data rate used by the slow node. This is because the reference throughput is constant for a given data rate, a given size payload and a given number of nodes. Simultaneously, TPA improves the aggregate throughput significantly. Specifically, when the low data rate is 1 Mbps, the aggregate throughput for TPA is more than twice that for DCF.

### Impact of traffic load

In this subsection, we tested TPA as a function of the number of nodes. Only one pair ofnodes is transmitted at 1 Mbps, and the other pairs at 11 Mbps.

Fig.1 shows that TPA achieves much higher throughput and better Jain fairness index than DCF. In each case, TPA can achieve the aggregate throughput, which almost equals the sum of the reference throughput of each flow.



(a)

(b)

**Fig.1 (a) Aggregate throughput versus the number of nodes; (b) Jain fairness index versus the number of nodes**

For the TPA scheme, the aggregate throughput increases as the traffic loads become heavier since the idle slots are reduced. After the number of nodes exceeds 12, the aggregate throughput starts to decrease because the number of collisions increases.

For the DCF scheme, the aggregate throughput keeps increasing as the traffic loads increase. This is because the throughput increment caused by the increasing number of flows transmitted at 11 Mbps exceeds the throughput decrement caused by the increasing number of collisions.

### Comparison between TPA and some other solutions

To further investigate the performance of the proposed scheme, we carried out simulations to compare the performance of TPA with some most relevant solutions:

1. Maximum transfer unit size adaptation approach (MTUA)

We adopt MTUA proposed by Yoo *et al.*(2005). The packet size is adjusted in the following way: $SMTU=LMTU\cdot r/11\text{e}6$, where $r$ represents the transmission rate of a node and *LMTU* represents the *MTU* of the nodes with 11-Mbps data rate.

2. Initial contention window adaptation approach (CWA)

Kim *et al.*(2005) proposed to resolve the performance anomaly by configuring the initial contention window size $CW_{\min}$ inversely proportional to the data rate. As mentioned earlier, this solution ignores the overhead. In order to make a fair comparison, we modified CWA by taking into account the overhead. The contention window size *CW* is adapted in the following way: $CW_{\min}(l)/CW_{\min}(k)=T_{\text{s}}(l)/T_{\text{s}}(k)$.

In this experiment, half flows are transmitted at 11 Mbps and the other half at 1 Mbps. For MTUA, the packet size for the fast node is set to 1000 bytes, while that for the slow node is set to 476 bytes. Table 3 shows the achieved throughput of each flow, the aggregate throughput, the collision rate, the Jain fairness index and the throughput gain for these three schemes. The throughput of each flow is the average throughput achieved by all the flows with the same data rate. The throughput gain is the gain on the average throughput of a new scheme (MTUA, CWA and TPA), compared with the basic DCF. It is calculated as: $Throughput\_gain=(S-S_{\text{DCF}})/S_{\text{DCF}}\times 100\%$.

As shown in Table 3, MTUA is less efficient than the other schemes because the overhead introduced by the header and backoffs leads to the degradation of the throughput achieved by the slow nodes. Only the fast nodes obtain the throughput close to the reference throughput.

The good fairness indexes and the high aggregate throughput for CWA and TPA indicate each flow achieves the needed throughput and the performance anomaly is solved cleanly. We can note that there are only small differences in each metric between CWA and TPA. CWA achieves slightly higher throughput than TPA while TPA has slightly better Jain fairness index than CWA. The fact indicates that TPA can provide the fairest channel occupancy time assignment.

**Table 3 Comparison between MTUA, CWA and TPA**

| Methods | Parameters | Node No. | | | | | |
|---|---|---|---|---|---|---|---|
| | | 4 | 8 | 12 | 16 | 20 | 24 |
| MTUA | 1 M Th. (kbps) | 126.653 | 58.597 | 39.477 | 29.131 | 23.468 | 19.191 |
| | 11 M Th. (kbps) | 2739.460 | 1567.940 | 996.955 | 716.255 | 544.796 | 439.973 |
| | Total Th. (kbps) | 2866.113 | 3253.075 | 3109.296 | 2981.546 | 2841.319 | 2754.985 |
| | Th. gain (%) | 99.9 | 140.3 | 146.7 | 144.5 | 148.4 | 143.1 |
| | Fairness index | 0.770 | 0.731 | 0.746 | 0.753 | 0.756 | 0.770 |
| | Collision rate (%) | 5.8 | 10.9 | 17.0 | 22.1 | 25.6 | 29.0 |
| CWA | 1 M Th. (kbps) | 363.929 | 177.254 | 117.529 | 86.394 | 68.129 | 54.137 |
| | 11 M Th. (kbps) | 2892.311 | 1492.395 | 965.873 | 698.432 | 537.450 | 442.460 |
| | Total Th. (kbps) | 3256.240 | 3339.297 | 3250.205 | 3139.305 | 3027.895 | 2979.586 |
| | Th. gain (%) | 127.1 | 146.7 | 157.9 | 157.5 | 164.7 | 162.9 |
| | Fairness index | 0.987 | 0.985 | 0.989 | 0.991 | 0.993 | 0.991 |
| | Collision rate (%) | 1.5 | 7.4 | 12.5 | 16.6 | 20.3 | 23.2 |
| TPA | 1 M Th. (kbps) | 367.138 | 183.541 | 122.181 | 89.113 | 68.525 | 58.873 |
| | 11 M Th. (kbps) | 2881.744 | 1454.207 | 935.870 | 686.257 | 539.962 | 420.789 |
| | Total Th. (kbps) | 3248.882 | 3275.497 | 3174.153 | 3101.478 | 3042.439 | 2877.967 |
| | Th. gain (%) | 126.6 | 142.0 | 151.8 | 154.4 | 165.9 | 153.9 |
| | Fairness index | 0.989 | 0.991 | 0.995 | 0.995 | 0.994 | 0.999 |
| | Collision rate (%) | 1.5 | 7.7 | 12.8 | 16.6 | 19.8 | 22.6 |

## CONCLUSION

IEEE 802.11 based wireless LANs has become more and more popular due to its simplicity and effectiveness. However, performance anomaly happens in the case that some nodes have lower transmission rates than others in the same wireless cell. Because 802.11 guarantees an equal long run channel access probability, high bit-rate nodes have the same throughput as the lowest bit-rate nodes so that the aggregated throughput degrades. To solve this problem, we propose a novel mechanism by introducing the transmission probability $p_t$ to 802.11b protocol. By configuring $p_t$ according to the bit-rate, the mechanism not only provides identical channel utilization to improve the throughput but also presents an effective service differentiation measure applying to multi-rate WLANs. Theoretical analysis and simulation results demonstrated that our scheme can eliminate the performance anomaly in many scenarios. The objective of time-based fairness and the needed aggregate throughput can be reached under each tested configuration.

### References

Bianchi, G., 2000. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE JSAC*, **18**(3):535-547. [doi:10.1109/49.840210]

Dunn, J., Neufeld, M., Sheth, A., Grunwald, D., Bennett, J., 2006. A practical cross-layer mechanism for fairness in 802.11 networks. *Mobile Networks and Applications*, **11**(1):37-46. [doi:10.1007/s11036-005-4459-z]

Garroppo, R.G., Giordano, S., Luceti, S., Tavanti, L., 2007. Providing air-time usage fairness in IEEE 802.11 networks with the deficit transmission time (DTT) scheduler. *Wireless Networks*, **13**(4):481-495. [doi:10.1007/s11276-006-9201-7]

Heusse, M., Rousseau, F., Berger-Sabbatel, G., Duda, A., 2003. Performance Anomaly of 802.11b. INFOCOM, p.836-843.

Iannone, L., Rousseau, F., Berger-Sabbatel, G., Duda, A., 2005. Idle Sense: An Optimal Access Method for High Throughput and Fairness in Rate Diverse Wireless Lans. SIGCOMM, p.121-132.

Jain, R., 1999. Throughput Fairness Index: An Explanation. ATM Forum/99-0045.

Kim, H., Yun, S., Kang, I., Bahk, S., 2005. Resolving 802.11 performance anomalies through QoS differentiation. *IEEE Commun. Lett.*, **9**(7):655-657. [doi:10.1109/LCOMM.2005.1461695]

Li, B.J., Soung, C.L., 2005. Proportional Fairness in Wireless LANs and Ad Hoc Networks. WCNC, **3**:1551-1556. [doi:10.1109/WCNC.2005.1424745]

Razafindralambo, T., Lannone, L., 2006. Dynamic Packet Aggregation to Solve Performance Anomaly in 802.11. MSWIM. Torremolinos, Malaga, Spain, p.247-254. [doi:10.1145/1164717.1164761]

Sadeghi, B., Kanodia, V., Sabharwal, A., Knightly, E., 2002. Opportunistic Media Access for Multi-rate Ad Hoc Networks. MOBICOM, p.486-497.

Tan, G., Guttag, J., 2004. Time-based Fairness Improves Performance in Multi-rate WLANs. USENIX, p.269-282.

Yoo, S., Choi, J., Hwang, J., Yoo, C., 2005. Eliminating the performance anomaly of 802.11b. *LNCS*, **3421**:1055-1062.

## APPENDIX: CALCULATING ACCESS PROBABILITY FOR PROPOSED PROTOCOL

We derive the access probability $\tau$ for the proposed protocol through Markov model under the saturation conditions. This paper extends the model in (Bianchi, 2000) by introducing a new parameter $p_t$ to the 802.11 protocol, as shown in Fig.A1.

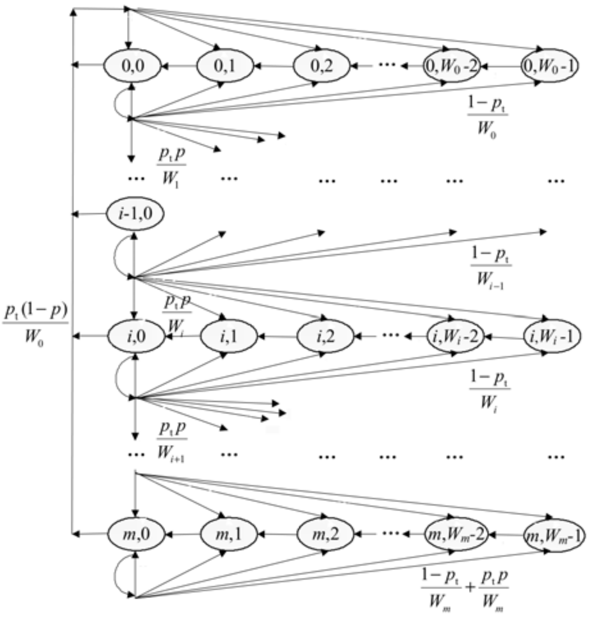


**Fig.A1 The Markov chain model for the proposed protocol**

Let $s(t)$ be the stochastic process representing the backoff stage of the node at time $t$ and $b(t)$ the stochastic process representing the backoff time counter. Let $m$ be the maximum backoff stage and $W$ the initial contention window. We adopt the short notation as (Bianchi, 2000): $P\{i_1, k_1|i_0, k_0\}=P\{s(t+1)=i_1, b(t+1)=$

$k_1|s(t)=i_0,\ b(t)=k_0\}$. Then the only null one-step transition probabilities are:

$$\begin{cases} P\{i,k\,|\,i,k+1\}=1, & i\in[0,m],\quad k\in[0,W_i-2],\\ P\{0,k\,|\,i,0\}=p_{\mathrm{t}}(1-p)/W, & i\in[1,m],\quad k\in[0,W-1],\\ P\{i,k\,|\,i,0\}=(1-p_{\mathrm{t}})/W_i, & i\in[1,m-1],\ k\in[0,W_i-1],\\ P\{i,k\,|\,i-1,0\}=p_{\mathrm{t}}p/W_i, & i\in[1,m],\quad k\in[0,W_i-1],\\ P\{0,k\,|\,0,0\}=p_{\mathrm{t}}(1-p)/W+(1-p_t)/W, & k\in[0,W-1],\\ P\{m,k\,|\,m,0\}=p_{\mathrm{t}}p/W_m+(1-p_t)/W_m, & k\in[0,W_m-1]. \end{cases} \tag{A1}$$

Let $b_{i,k}=\lim_{t\to\infty}P\{s(t)=i,\ b(t)=k\}$, $i\in[0,\ m]$, $k\in[0,\ W_i-1]$ be the stationary distribution of the chain. It is easy to obtain the following equations:

$$\begin{cases} b_{i,0}=p_{\mathrm{t}}pb_{i-1,0}+(1-p_{\mathrm{t}})b_i\\ \qquad\Rightarrow b_{i,0}=p^ib_{0,0}, \qquad i\in[0,m-1],\\ b_{m,0}=p_{\mathrm{t}}pb_{m-1,0}+p_{\mathrm{t}}pb_{m,0}+(1-p_{\mathrm{t}})b_{m,0}\\ \qquad\Rightarrow b_{m,0}=\dfrac{p^m}{1-p}b_{0,0}. \end{cases} \tag{A2}$$

Noting that $\sum_{i=0}^{m}b_{i,0}=b_{0,0}/(1-p)$, for $i\in[1,m-1]$, we have:

$$\begin{cases} b_{0,W-1}=\dfrac{p_{\mathrm{t}}(1-p)}{W}\sum_{i=0}^{m}b_{i,0}+\dfrac{1-p_{\mathrm{t}}}{W}b_{0,0}=\dfrac{b_{0,0}}{W},\\ b_{m,W_m-1}=\dfrac{p_{\mathrm{t}}p}{W_m}b_{m-1,0}+\dfrac{1-p_{\mathrm{t}}}{W_m}b_{m,0}+\dfrac{p_{\mathrm{t}}p}{W_m}b_{m,0}\\ \qquad=\dfrac{p^m}{W_m(1-p)}b_{0,0},\\ b_{i,W_i-1}=\dfrac{p_{\mathrm{t}}p}{W_i}b_{i-1,0}+\dfrac{1-p_{\mathrm{t}}}{W_i}b_{i,0}=\dfrac{p^i}{W_i}b_{0,0}. \end{cases} \tag{A3}$$

From Eqs.(A2) and (A3), when $k\in[0,\ W_i-2]$, $i\in[1,m-1]$, $b_{i,k}$ can be expressed as functions of $b_{0,0}$ and the packet collision probability $p$:

$$\begin{cases} b_{0,k}=\dfrac{p_{\mathrm{t}}(1-p)}{W}\sum_{i=0}^{m}b_{i,0}+\dfrac{1-p_{\mathrm{t}}}{W}b_{0,0}+b_{0,k+1}\\ \qquad=b_{0,W-1}+b_{0,k+1}=\dfrac{W-k}{W}b_{0,0},\\ b_{m,k}=\dfrac{p_{\mathrm{t}}p}{W_m}b_{m-1,0}+\dfrac{1-p_{\mathrm{t}}}{W_m}b_{m,0}+\dfrac{p_{\mathrm{t}}p}{W_m}b_{m,0}+b_{m,k+1}\\ \qquad=b_{m,W_m-1}+b_{m,k+1}=\dfrac{W_m-k}{W_m}\dfrac{p^m}{1-p}b_{0,0},\\ b_{i,k}=\dfrac{p_{\mathrm{t}}p}{W_i}b_{i-1,0}+\dfrac{1-p_{\mathrm{t}}}{W_i}b_{i,0}+b_{i,k+1}=b_{i,W_i-1}+b_{i,k+1}\\ \qquad=\dfrac{W_i-k}{W_i}p^ib_{0,0}. \end{cases} \tag{A4}$$

By using the normalization condition for stationary distribution, we obtain the solution of $b_{0,0}$:

$$\begin{aligned} 1&=\sum_{i=0}^{m}\sum_{k=0}^{W_i-1}b_{i,k}\\ &=b_{0,0}\sum_{i=0}^{m-1}\sum_{k=0}^{W_i-1}\frac{W_i-k}{W_i}p^i+b_{0,0}\sum_{k=0}^{W_m-1}\frac{W_m-k}{W_m}\frac{p^m}{1-p}\\ \Rightarrow b_{0,0}&=\frac{2(1-2p)(1-p)}{(1-2p)(W+1)+pW[1-(2p)^m]}. \end{aligned} \tag{A5}$$

When the backoff timer reaches zero, any station initiates a transmission at the probability $p_{\mathrm{t}}$. So the access probability $\tau$ is:

$$\tau=p_{\mathrm{t}}\sum_{i=0}^{m}b_{i,0}=\frac{p_{\mathrm{t}}b_{0,0}}{1-p}=\frac{2p_{\mathrm{t}}}{1+W+pW\sum_{i=0}^{m-1}(2p)^i}. \tag{A6}$$

Compared with the access probability $\tau'$ for the IEEE 802.11 given in (Bianchi, 2000), the relation between $\tau$ and $\tau'$ is:

$$\tau=p_{\mathrm{t}}\tau'. \tag{A7}$$