Delay Analysis of an Enhancing IEEE 802.11 Point Coordination Function MAC Protocol

Zheng Guan, Hongwei Ding, and Wenhua Qian
School of Information Science & Technology, Yunnan University, Kunming Yunnan, China
Email: gz_627@sina.com, dinghw@ynu.edu.cn, qwhua003@sina.com

Abstract—The widespread use of multimedia networking applications has brought more requirements to the network, creating a need for end-to-end Quality of Service (QoS). The contention-free Point Coordination Function (PCF) with Round Robin scheme, where the Access Point controls all transmissions based on a polling mechanism, is defined in IEEE 802.11 standard to provide QoS service. This network suffers from inefficiency in the delay sensitive traffic and makes. It is difficult to provide priority-differentiated service. In this paper, a useful enhance mechanism is derived from a two-level-polling model with a parallel scheme between the polling and switch over processes. We consider two different polling schemes according to station classify: key station and normal station. This mechanism allows for a delay guarantee as well as priority service. An analytical is developed for evaluating the queuing length for the station and waiting time for the packets at nodes. Based on this model, we achieve several closed-form expressions for mean cycle time, mean queue length and mean waiting time. To verify the correctness of our analytical model, we also develop a simulator for the 802.11 PCF MAC. The simulation results well match the analytical results.

Index Terms—IEEE802.11; PCF; Delay Guarantee; Priority Service; Mean Waiting Time

I. INTRODUCTION

With the large-scale deployments of wireless local area networks (WLANs) in homes, offices, and public areas, the IEEE 802.11 has become one of the dominant technologies in broadband wireless network communication. The point coordination function (PCF) and the distribution coordination function (DCF) with round robin scheduling are two foundational control mechanism in media accesses [1]. However, the former has the drawbacks of low utilization of resource and the latter suffers from collision seriously under heavy load. As many new applications for 802.11 enabled devices emerge which require higher quality of services and better mobility support, there is an increasing demand for improving the performance of 802.11 networks. As discussed in [2] and [3], the widespread use of multimedia networking applications brought more requirements to the network. Many new applications for 802.11 enabled devices emerge which require higher quality of services, and then increasing the demand for end-to-end quality of service in the network [4].

We have observed a number of works dealing with the PCF, it related to improve the overall network performance through novel scheduling algorithms like [5] and [6], designing new polling mechanisms to reduce the overhead associated to the polling process, and extending PCF to networks without infrastructure as shown from [7] to [10]. Refraining from polling overhead to improve utilization and the delay performance, providing differentiation service are the key aims. Biplab analytically characterizes the delays experienced with the IEEE 802.11 PCF where nodes may employ the power management modes specified in the standards in order to conserve energy in [11]. Kuan-Hung considered two different polling schemes: poll once and multiple poll for packets transmission according to packet aggregation [12]. However, it seems difficult to meet the demands of the priority service and real-time applications for the key station. There are few efforts in QoS guarantee and the low delay satisfaction of the sensitive traffic from priority service. The exception can be found in [13][14]and [15] where a two-level polling model are created into MAC protocol in order to realize the differentiation for the key station with real-time and priority service. In [14], a second level of deterministic channel access on top of the traditional IEEE 802.11 DCF protocol is introduced to grant higher priority to high-rate stations in a fully distributed manner, and collisions among potential contending stations. In [13], Bao proposed a continuous time scheduling algorithm with two level priorities based on clustering. The high-priority service of the cluster head is responsible for communication between the different low-priority service cluster nodes. Ref. [14] expanded this model to the discrete-time area, and applied in the IEEE802.11 PCF to enhance the performance of RR scheme. However, it seems a waste of time as handoff time is needed between key station and normal station in the two-level algorithm.

Motivated by this, we propose an improving PCF model. Our improving PCF works at a two-level polling model, similar to other existing two-level polling mechanisms in [10] and [13]. However, there are several key differences: Data transmit and request could be processed parallel. When the base station (AP) polls the non-empty node, the polling message from the AP can be piggybacked with the ACK packet; also, the queue state of a node can be piggybacked with the data packet. Therefore, the polling scheme synchronizes the polling and transmission of information packets no additional overhead.
Using theoretical analysis coupled with extensive simulations, we show that the delay performance as well as the system stability is improved.

II. SYSTEM MODEL

A. System Working Condition

In order to focus on the MAC layer, all the packets are assumed to be received without errors and thus the results herein presented correspond to an upper-bound of the performance of the protocol.

$N+1$ stations in a WLAN are categorized as one key station, noted by $h$, with high priority real-time data service and $N$ normal stations, noted by $i = 1, 2, \ldots, N$. Station has two modes: polling mode and idle mode.

A station is an idle station if it has no data to send. Non-empty station in polling mode waits for polling messages from the AP by checking certain header fields in the ACK. If polled, it will send a packet without backing off for any time slot. For the last packet in the queue, if ACK is received, it will leave for the idle mode immediately until receive new data from applications.

At this stage, a two-level polling scheme is used to provide a priority service and a parallel mechanism will enhance the PCF delay-guaranteed performance.

Data access is started by the receiver. AP starts the access process as the controller by sending a data request to the stations according to a polling list, in which the key station is assigned between two successive normal stations.

Furthermore, the data transmit and data request are process parallel by the piggyback technical. When the AP receives the last data packet with $I_{\text{last}}=1$ form a certain station, it will send an ACK with a data request message to the sender, in which $S_{\text{req}}$ is set as the number of the follower in the polling list. At this time, if the successive station works in the polling mode, it is listening to the channel and checks this data request message, then begin sending data packets immediately. If the successive station works in the idle mode, there will be no responding, when the AP timed out waiting for the request, it will send a data request packet to the next station in the polling list. Following is the packet format.

The normal station works in a RR (Round Robin) mechanism with 1-limited service, and continuously monitor the data channel, whenever a key station is detected on the data channel, the normal must yield to the key station.

B. System Model Discription

Most of the existing performance analysis method of PCF is tedious and complicated. Ref. [17] applied the classic M/G/1 vacation model to analyze the delay performance of PCF with one-level polling. In this section we present the proposed model [16] to evaluate the delays experienced by nodes using this enhancing PCF mode with two-level parallel polling.

Assume a random variable $\xi_i(n)$ ($i = 1, 2, \ldots, N$) as number of data packets in storage at normal station $i$ at time $t_n$. The base station (AP) polls normal station $i$ at time $t_n$ switches to poll key station $h$ at time $t_i^*$ and then polls normal station $i+1$ at time $t_{i+1}$ ($t_n < t_i^* < t_{i+1}$). Then the status of the entire polling model can be represented as $\{\xi_1(n), \xi_2(n), \ldots, \xi_n(n), \xi_h(n)\}$.

$\{\xi_1(n+1), \xi_2(n+1), \ldots, \xi_n(n+1), \xi_h(n+1)\}$ and $\{\xi_1(n'), \xi_2(n'), \ldots, \xi_n(n'), \xi_h(n')\}$. For $\sum_{i=1}^N \rho_i + \rho_h < 1$ ($\rho_i = \lambda_i\beta_i$, $\rho_h = \lambda_h\beta_h$), we will always assume the stations are stable. Then, the probability distribution is defined as

$$\lim_{n\to\infty} P[\xi_j(n) = x_j; j = 1, 2, \ldots, N, h]$$

$$= \pi_j(x_1, \ldots, x_j, \ldots, x_N)$$

By using the embedded Markov chain theory and the probability generating function method to characterize the proposed system model, the generating function of the normal station at the polling time is the following.

$$G_i(z_1, \ldots, z_i, \ldots, z_N, z_h) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \sum_{x_N=0}^{\infty} \sum_{x_h=0}^{\infty} z_1^{x_1} \cdots z_i^{x_i} \cdots z_N^{x_N} \pi_i(x_1, \ldots, x_i, \ldots, x_N, x_h)$$

(1)

According to the proposed mechanism, the system variables have the following equations.
\[ \xi_j(n') = \xi_j(n) + \eta_j(v_j), \quad j \neq i, \quad \xi_i(n) = 0 \]
\[ \xi_j(n') = \xi_j(n) + \eta_j(v_j) - 1, \quad j = i, \quad \xi_i(n) = 0 \]
\[ \xi_i(n') = \xi_i(n) + \mu_i(u_i), \quad j \neq i, \quad \xi_j(n) = 0 \]
\[ \xi_j(n+1) = \xi_j(n') + \eta_j(v_j), \quad j \neq h \]
\[ \xi_j(n+1) = 0, \quad j = h \]

\( v_j(n) \) is the service time in station \( j \), \( \eta_j(v_j) \) is the number of arrivals to station \( k \) during service duration of \( Q_i \), \( u_i(n) \) is the switchover time from \( Q_i \) to node \( Q_k \) when station \( i \) is an idle station, and \( \mu_i(u_i) \) is the number of arrivals to \( Q_i \) during \( u_k(n) \) (\( j = 1,2,\ldots,N,h ; \ k = 1,2,\ldots,N,h \)).

Then the generating function of the key station \( Q_k \) at the time \( t_n \) is the following.

\[
G_{\lambda_k}(z_1,\ldots,z_i,\ldots,z_N,z_k) = \lim_{n \to \infty} \mathbb{E} \left[ \prod_{j=1}^{N} \xi_j^{(n')} z_j^{(n')} \right]
\]
\[= \frac{1}{z_k} B_k \left( A_k(z_k) \prod_{j \neq k} A_j(z_j) \right) \left[ G_k(z_1,\ldots,z_i,\ldots,z_N,z_k) \right]_{z=0}
- G_k(z_1,\ldots,z_i,\ldots,z_N,z_k)\big|_{z=0}
+ R_k \left( A_k(z_k) \prod_{j \neq k} A_j(z_j) \right) \cdot G_k(z_1,\ldots,z_i,\ldots,z_N,z_k) \big|_{z=0}
\]

where \( G_k(z_1,\ldots,z_i,\ldots,z_N,z_k) \) is the generation function for the number of data packets present at polling instants of \( Q_i \). \( A_j(z_j) \), \( B_j(z_j) \), and \( R_j(z_j) \) represent the generation functions of the arrival, service and switch over procedure in \( Q_i \), \( j = 1,2,\ldots,N,h \).

The generating function for the number of data packets present at polling instants \( t_{n+1} \) is

\[
G_{\lambda_{n+1}}(z_1,\ldots,z_i,\ldots,z_N,z_k) = \lim_{n \to \infty} \mathbb{E} \left[ \prod_{j=1}^{N} \xi_j^{(n+1')} z_j^{(n+1')} \right]
\]
\[
= G_k \left[ z_1,\ldots,z_i,\ldots,z_N,R_k \left( \prod_{j \neq k} A_j(z_j) F_k \left( \prod_{j \neq k} A_j(z_j) \right) \right) \right]
\]

where \( F_k(z_k) \) is the generation function of the timing variables to serve data packets arrive at key station \( h \) in each time slot exhaustive.

### III. Performance Analysis

Let the average number of message packets at \( Q_i \) at \( t_s \) be defined as \( g_i(j) \) when \( Q_i \) is polled and at \( t_{n+1} \) as \( g_{\lambda_k}(j) \) when \( Q_k \) is polled. Then \( g_i(j) \) and \( g_{\lambda_k}(j) \) can be given as the following.

\[
g_i(j) = \lim_{z \to z_1,\ldots,z_i,\ldots,z_N,\lambda_i \to \infty} \frac{\partial G_k(z_1,\ldots,z_i,\ldots,z_N,z_k)}{\partial \lambda_i}
\]
\[i = 1,2,\ldots,N; \; j = 1,2,\ldots,N,h \]

\[
g_{\lambda_k}(j) = \lim_{z \to z_1,\ldots,z_i,\ldots,z_N,\lambda_k \to \infty} \frac{\partial G_k(z_1,\ldots,z_i,\ldots,z_N,z_k)}{\partial \lambda_k}
\]
\[i = 1,2,\ldots,N; \; j = 1,2,\ldots,N,h \]

And let

\[
g_i(j,k) = \lim_{z \to z_1,\ldots,z_i,\ldots,z_N,\lambda_i \to \infty} \frac{\partial^2 G_k(z_1,\ldots,z_i,\ldots,z_N,z_k)}{\partial \lambda_i \partial \lambda_k}
\]
\[i = 1,\ldots,N; \; j,k = 1,\ldots,N,h \]

\[
g_{\lambda_k}(j,k) = \lim_{z \to z_1,\ldots,z_i,\ldots,z_N,\lambda_k \to \infty} \frac{\partial^2 G_k(z_1,\ldots,z_i,\ldots,z_N,z_k)}{\partial \lambda_i \partial \lambda_k}
\]
\[i = 1,\ldots,N; \; j,k = 1,\ldots,N,h \]

#### A. Mean Cycle Time

The mean cyclic period, \( \bar{\theta} \), is the mean value of the time between two successive visit beginnings to station \( i \); it consists of service time and switchover time. In the proposed mechanism, it is zero-switchover time system when the buffer is not empty, and then decrease the cycle time. Considering the characteristics of the generating function, the related expressions can be given as Eq. (3).

\[
1 - G_k(z_1,\ldots,z_i,\ldots,z_N,z_k) \big|_{z=0} = \lambda \bar{\theta} \quad j = 1,2,\ldots,N,h
\]

Take Eq. (3) and Eq. (4) into Eq. (5) and Eq. (6), simplify these using Eq. (9), we have:

\[
\bar{\theta} = \frac{1 - G_k(z_1,\ldots,z_i,\ldots,z_N,z_k) }{\lambda_i}
= \frac{N\gamma}{1 - \rho_\lambda - N\rho + N\lambda\gamma}
\]

#### B. Mean Queue Length

According to the mechanism of exhaustive service and 1-limited service, the mean queue length of the key station can be derived from substituting Eq. (3) and (4) into Eq. (5) and Eq. (6), respectively, and \( g_{\lambda_k}(h) \) can be expressed as the following.

\[
g_{\lambda_k}(h) = \lambda_k \left[ \gamma + \lambda (\beta - \gamma) \bar{\theta} \right]
\]

The mean queue length of the normal station can be derived from substituting Eq. (3) and Eq. (4) into Eq. (7) and Eq.(8), considering the special characteristics of the symmetric polling system as well as the set of discrete-time equations, then \( g_i(j) \) can be expressed as the following:
The mean queue length for the center queue within interval service can be derived from substituting Eq. (4) into Eq. (8), and then \( g_a(i) \) can be expressed as the Eq. (13).

\[
g_a(i) = \frac{1 - \rho_n}{2(1 - N \rho - \rho_n)} \left( N \left( \gamma A(1) + \lambda^2 R(1) \right) + \frac{2N \rho_n}{1 - \rho_n} \left( \lambda^2 R(1) + \lambda^2 \right) + \frac{N \rho_n^2}{(1 - \rho_n)} \left( \gamma A(1) + \lambda^2 R(1) \right) \right) + \frac{N \lambda^2}{1 - N \rho - \rho_n + N \lambda N} \left( N \gamma^2 \beta(1) + \lambda^2 \gamma R(1) \right) + \frac{1}{1 - \rho_n} \left( 2N \lambda^2 \rho_n(\beta - \gamma) - 2N \lambda^2 \gamma (1 + \rho_n) R(1) \right)
+ 2(1 - \rho - \rho_n) + \lambda \rho_n (1 + \rho_n) - (N - 1) \lambda \gamma^2 + 2(N - 1) \lambda \gamma (\beta - \gamma) + \left( \frac{\lambda}{(1 - \rho_n)} \times N \lambda \gamma B(1) \right)

\]

\[
\text{C. Mean Waiting Time}
\]

The waiting time of message packets, \( w_n \), denotes the time from when a data arrives at the station \( j \) \((j = 1, 2, \ldots, N, h)\) from applications to when it is served, and \( E[w_n] \) as well as \( E[w_h] \) denotes the mean waiting time of packets for the normal station and the key station, respectively, for normal stations [11]:

\[
E[w_i] = \frac{1}{A} g_a(i) - \frac{1}{\lambda} A(1) \gamma
\]

Take with Eq. (12), it is easy to obtain the closed form expression of mean waiting time.

The key station is served in the exhaustive scheme, under the theory in [11], we could get the expression of the mean waiting time as follow:

\[
E[w_k] = \frac{g_a(h, h)}{2 \lambda^2 g_a(h, h)} - \frac{A_1(h)}{2 \lambda^2} + \frac{A_1(h)}{2(1 - \rho_n)}
\]

\[
IV. \text{NUMERICAL RESULT}
\]

Research in [13] has demonstrated the enhanced PCF with two level mixed service polling scheme has lower mean waiting time than classical RR mechanism. In this section, we illustrate the comparisons of the analytical and simulation results between the two level mixed service PCF and the improving scheme presented in this letter.

\[
\text{A. Assumption and the Basic Parameter Values}
\]

Consider a polling system with ten queues-one key station and nine normal stations. The system parameters are shown in Table. I. All the packets are assumed to be received without errors. In this example we illustrate the accuracy of the theoretical analysis and simulation results of the model performance under the mechanism of both schemes.

The simulation was run for at least 1,000,000 time units and sometimes longer for higher utilizations. Basic parameter values

\[
\text{TABLE I. BASIC PARAMETER VALUES}
\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Tx. Rate</td>
<td>54Mbps</td>
<td>Slot Time</td>
<td>10μs</td>
</tr>
<tr>
<td>Data Packet Length</td>
<td>1250 bytes</td>
<td>Waiting-timeout</td>
<td>20μs</td>
</tr>
<tr>
<td>CTS/ACK packets</td>
<td>14bytes</td>
<td>RTS/Poll packet</td>
<td>20bytes</td>
</tr>
</tbody>
</table>

From the above comparisons, some results for the present schemes can be remarked as follows.

When the arrival rate of message packets is increased, the MQL as well as the MWT of the common station is gradually increased, while the increasing in key station is much smooth. Furthermore, the MWT of both the normal station and the key station are lower than the previous one, and its performance seems to be affected by the growth of the arrival rate slightly, this indicates that comparing with the previous, the proposed scheme can guarantee low delay along the real time service for both the key station and the common station.

As for the flow balance condition, the previous scheme is under the condition of \( N \rho + \rho_n + N \lambda \gamma < 1 \), whereas the proposed scheme is under the condition of \( N \rho + \rho_n < 1 \), so the latter has less constrains than that of...
the former. When $\lambda$, $\lambda_n$, or $N$ increased, the latter has greater stable working rang, when $\lambda$, $\lambda_n$, or $N$ exceeds the corresponding outbound, the common station will suffer in the saturation, and the MQL as well as the MWT will increase significantly, as shown in Fig. 4 and Fig. 6.

![Figure 3. Mean queue length of key station](image)

![Figure 4. Mean queue length of normal station](image)

![Figure 5. Mean waiting time of key station](image)

![Figure 6. Mean waiting time of normal station](image)

V. CONCLUSION

This paper has presented an enhancing IEEE 802.11 PCF mechanism. Performance evaluation and simulation results clearly show that the presented mechanism improves QoS support in IEEE802.11 networks, in special two-level polling and parallel processing are used.

Like RR, enhanced PCF in this letter includes a polling mechanism controlled by AP. Nevertheless it allows QoS differentiation because in the two level polling route and different service policy, the key terminator could be assigned more resources, which is an important improvement over RR. Moreover, compares to other existing two-level-polling mechanisms, the present scheme synchronizes the polling and transmission of information packets no additional overhead. Due to the piggyback technique, it presents lower transmission delays as the switch over time decreases and in special for traffic with heavy load.

At last, the mathematical model is developed to analyze the proposed scheme, and the closed form expressions of the MQL and MWT are derived. The efficiency of this mechanism is evaluated by mathematic analysis and proved by simulation results.

REFERENCE


Zheng Guan births in China, June, 1982. Received her M.Sc. degree in Signal and Information Processing from Yunnan University, China, in 2007, and the Ph.D. degree in Communication and information Systems, at the same university.

Her main research interests are in polling systems and their applications, wireless network access control protocol. She is a lecturer in School of information science and technology, Yunnan University, China

Hongwei Ding births in China, 1964. Received his Ph.D. degree in Communication and information Systems from Yunnan University. Now he is an associate professor in the same university. His current research interests include wireless sensor network.

Wenhua Qian births in China, January, 1980. Received his M.Sc. (2005) and Ph.D. in Computer Sciences (2010) from Yunnan University. Now he is lecturer at Computer Science and Engineering Department of Yunnan University. His current research interests include network communication. He has co-authored more than 30 papers.