Delay Model for Super-Frame Based Resource Reservation in Distributed Wireless Networks

Xiaobo Yu, Pirabakaran Navaratnam, and Klaus Moessner

University of Surrey, Guildford, GU2 7XH, UK
{x.yu,p.navaratnam,k.moessner}@surrey.ac.uk

Abstract. This paper proposes an analytical framework for evaluating the delay performance of super-frame (SF) based MAC schemes with distributed resource reservation in IEEE 802.11e enhanced distributed channel access (EDCA). SF-based resource reservation (RR) schemes divide the airtime into service intervals (SIs) with contention-free period (CFP) for providing guaranteed QoS for RTSNs and contention access period (CAP) for pledging fairness toward other sessions. The proposed analytical framework models the delay performance of RTSNs that obtain dedicated resources in a distributed manner. In addition, the optimization of system parameters, such as size of transmission opportunity (TXOP) and SI are studied in order to enhance the overall network capacity. The accuracy of the analytical framework is verified through numerical simulation and analytical results, which also suggest that the optimum resource allocation and SI can be found for improving the network capacity.

Keywords: QoS, IEEE 802.11e, resource reservation.

1 Introduction

Nowadays, IEEE 802.11-based wireless communication technology pervades in various areas such as Wi-Fi hot spots, city wide mesh networks, vehicular communication, and similar application areas. Most personal communication devices such as laptop computers as well as mobile phones are armed with 802.11a/b/g adapters or 802.11-compliant entities. Despite of the general application, there are still lots of issues that pose difficulties in providing Quality of Service (QoS) in 802.11-based distributed wireless networks.

So far, many research works have been focused on providing QoS for real-time sessions (RTSNs) in IEEE 802.11-based distributed wireless networks. Since the legacy distributed coordination function (DCF) can not differentiate the services between RTSNs and non-real-time sessions (NRTSNs), an enhancement of DCF named enhanced distributed channel access (EDCA) has been standardised in IEEE 802.11e [1]. Its fundamental QoS support is proved helpful for QoS support but its enhancement is still limited. For further improving the QoS for EDCA, some of the contributions [2, 3] are made to enhance the probability of channel access for RTSNs by tuning the parameters of deferral and back-off algorithms.

L. Atzori, J. Delgado, and D. Giusto (Eds.): MOBIMEDIA 2011, LNICST 79, pp. 89–104, 2012.

Optimizations of queueing algorithm [4–6] for improving the QoS for RTSNs are also achieved. Although QoS advancements toward RTSNs can be implemented by using these approaches, the transmissions of RTSNs are still affected under interference environment and the deferral and back-off, as the channel access overheads, are unavoidable.

To solve the aforementioned issues, one of the most effective solutions is the super-frame (SF) based MAC scheduling mechanisms with resource reservation (RR), which can partition the channel airtime into contention-free period (CFP) for providing guaranteed QoS for RTSNs and contention access period (CAP) for the fairness toward other types of traffic sessions. The RTSNs can get periodic and dedicated resources through this distributed RR scheme so that their QoS requirements can be met. Following the idea of this distributed SF-based RR method, several MAC protocols [7, 8] have been proposed and simulations have been conducted for validating the effectiveness of these schemes. However, the analysis as well as the optimization of these distributed SF-based RR schemes are still open issues.

This paper mainly proposes an analytical framework for modelling the delay performance of QoS guaranteed RTSNs in the distributed SF-based mechanisms ,that were devised for IEEE 802.11e EDCA. The analytical framework is capable of predicting QoS performance of RTSNs on both saturated and unsaturated traffic conditions. Based on the guaranteed QoS, the enhancement of network capacity (i.e. the maximum amount of RTSNs) is studied through the optimization of bandwidth allocation for RTSNs as well as the system parameter such as service interval (SI) in order to accommodate more RTSNs in CFP. Note that the network capacity in this paper implies the maximum amount of RTSNs that are allowed to reserve transmission opportunities (TXOPs) in CFP.

The rest of this paper is organized as follows. Section 2 depict the SF based RR mechanism and its derivative protocols - EDCA/RR and EDCA/DRR. Section 3 specifies the analytical model for delay performance. The optimization study is presented in Section 4. Simulation and analytical outcomes are shown in Section 5. Finally, section 6 concludes this paper.

2 Overview of Super-Frame Based Resource Reservation in IEEE 802.11e Networks and Its Derivative Protocols

As mentioned before, the SF-based RR schemes utilize SI to partition the services between admitted RTSNs and other sessions. Fig. 1 shows an example of the SF-based RR scheduling, a QoS guaranteed RTSN will obtain a dedicated bandwidth called transmission opportunity (TXOP) during which multiple frames of the corresponding RTSN can get transmitted provided that the dedicated duration is adequate. If the residual time can not afford a further data transmission, the corresponding RTSN will wait until the next TXOP.

To implement this distributed RR, EDCA/RR [7] proposes a signalling process. An add traffic stream (ADDTS) request frame is broadcasted from the source of the RTSN if RR is required. The signalling frame takes the traffic

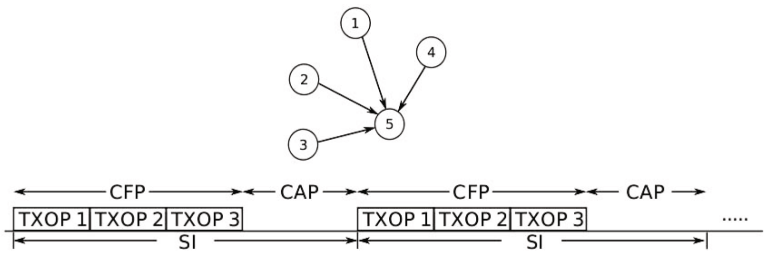

Fig. 1. Super-frame based resource reservation scheduling

specification (TSPEC) which contains the parameters such as service start time, delay bound, etc. Upon the receipt of the request frame, destination will decide whether to accept the RTSN and reserve bandwidth for it. If the residual bandwidth in CFP can support the QoS demand of the RTSN, destination will confirm the reservation request through replying an ADDTS response frame. Otherwise, it will send back the response frame to reject the RTSN. The reservation request is also validated by contending nodes within the transmission range. They will confirm the new request given that their dedicated resources are not offended. Otherwise, they decide to reject the new RTSN and inform the source by sending signalling messages.

Although EDCA/RR is able to successfully implement the SF-based RR, it ignores the dynamic resource allocation for the TXOPs that become idle after their corresponding RTSNs stop transmitting. This will incur the wastage of bandwidth in CFP and degrade the network capacity. In EDCA/DRR [8], a dynamic resource allocation scheme is proposed for addressing this problem. Arrival priority (APR) is introduced for differentiating the precedence of the rejected RTSNs that are made to be transmitted in CAP. The rule is that the earlier the RTSN accesses the CAP, the higher its APR is. An adaptive admission control is devised for monitoring and controlling the transmissions in CFP. If there are RTSNs being transmitted in CAP when idle resources appear in CFP, the idle resources will be assigned to the RTSN with the highest APR and then all the other rejected RTSNs shift their priority accordingly. If no rejected RTSN exists, the idle resources will be allocated to the CAP.

3 Analytical Model

In this section, an analytical framework for modelling the delay performance of RTSNs with dedicated resources is proposed for distributed SF-based RR mechanisms. To enhance the efficiency of transmission time and reduce the channel deferral time, MAC service data unit (MSDU) is formed by aggregating several frames of a session [9]. This can help improve the throughput of the session. During CFP, the amount of MSDUs for RTSN_i that are permitted to be transmitted within its dedicated TXOP depends on the duration of SI denoted by ΔSI, mean

MSDU size $\overline{s_{DATA_i}}$, and required scheduling rate which is represented by λ_i[1]. Therefore, we can obtain

$$n_{t,i} = \lceil \frac{\Delta SI \times \lambda_i}{\overline{s_{DATA_i}}} \rceil \tag{1}$$

where $n_{t,i}$ is the amount of MSDUs of RTSN$_i$ that are able to be accommodated by a TXOP. Fig. 2 shows an example of a scheduled TXOP which can be used for multiple data transmissions. The duration of a TXOP is expressed by

$$t_{TXOP_i} = n_{t,i}(E[t_{DATA_i}] + t_{ACK}) + 2n_{t,i} \cdot t_{SIFS} \tag{2}$$

where t_{ACK} and t_{SIFS} denote the duration of ACK and SIFS, respectively. $E[t_{DATA_i}]$ in the above equation stands for the average duration cost by transmitting an MSDU of RTSN$_i$. The analytical model for delay performance of QoS guaranteed RTSNs is specified as follows.

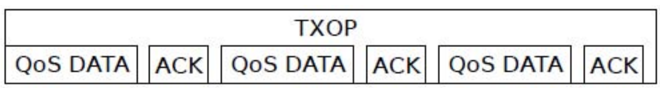

Fig. 2. A scheduled TXOP

3.1 Delay Model for RTSNs with TXOPs

In this subsection, we analyse the delay performance of RTSNs with TXOPs in CFP. The average delay $d_{ave,i}$ of MSDUs for RTSN$_i$ is equivalent to the average duration from the instant that its MSDU buffers in the queue to the moment that it successfully completes transmission. In general, the delay is comprised of channel access delay $d_{ca,i}$, queueing delay $d_{q,i}$ as well as transmission delay $d_{tr,i}$. The detail analyses are shown below.

Channel Access Delay. The channel access delay is defined as the time from an MSDU reaches the head of the interface queue to the instant that it starts accessing the channel. Owning dedicated bandwidth in CFP, each MSDU needs to wait for its time-slots in CFP to get transmitted. The instant when the MSDU arrives the head of queue determines its channel access delay. As shown in Fig. 3, for each RTSN, time can be regarded to be composed of periodic TXOP for data transmission and non-TXOP time during which its MSDUs have to wait. Based on the relationship between traffic load of a RTSN and the size of its allotted TXOP. Three conditions can be defined: (i). Unsaturated condition which indicates that the allocated resources can not be used entirely by the RTSN. (ii). Saturated condition which implies that the RTSN can exactly feed

[1] In this paper, it is assumed that application rate is equivalent to required scheduling rate.

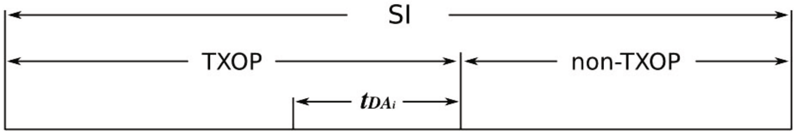

Fig. 3. TXOP, t_{DA_i} and non-TXOP

all the duration of TXOP. (iii). Over-saturated condition which indicates that the maximum transmission capability of reserved bandwidth is not sufficient for accommodating the traffic load of the RTSN. The third condition results in buffer overflow and thus devastates the performance of RTSN. Since the analytical model is aimed at evaluating the delay performance of QoS guaranteed RTSNs, the unsaturated and saturated conditions will be the focus hereafter.

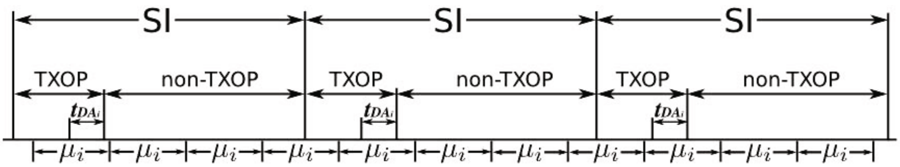

Fig. 4. SI, sending interval and TXOP

As a key parameter for analysing the channel access delay, the MSDU arrival time is mainly affected by the required scheduling rate λ_i of RTSN$_i$ and the duration of allocated TXOP t_{TXOP_i}. To formulate the channel access delay using the above parameters, sending interval is introduced to map the required scheduling rate into the periodic SI. Let μ_i denote the sending interval of RTSN$_i$. We can obtain

$$\mu_i = \frac{\overline{s_{DATA_i}}}{\lambda_i} \tag{3}$$

As shown in Fig. 4, the relationship between the required scheduling rate and the TXOP can be easily indicated if sending interval is employed. Assumed that the MSDUs of RTSNs regularly generate, it is able to figure out whether the instant that an MSDU arrives at the queue belongs to its TXOP or its non-TXOP. Since the channel access delay is determined by the arrival instant of each MSDU, a function $f = \delta(x)$ is given for representing normalized offset of each arrival within a SI. It can be expressed by

$$\delta(x) = x - [x] \tag{4}$$

where [x] is used for taking the integer part of variable x. An example of a normalized offset of MSDU arrival time is shown in Fig. 5. To express the offset duration of each arrived MSDU within its SI, $\eta(j)$ is introduced and expressed by

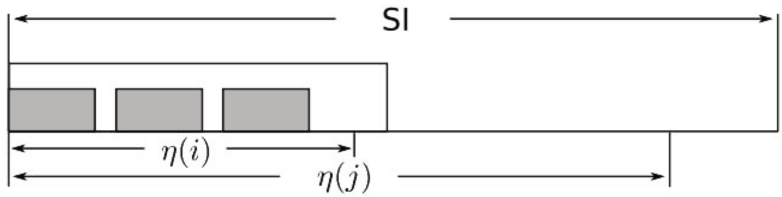

Fig. 5. Arrival offset

$$\eta(j) = \delta(\frac{j \cdot \mu_i}{\Delta SI}) \cdot \Delta SI \tag{5}$$

where j stands for the jth MSDU of the RTSN. If a head-of-line MSDU arrives and finds out that the residual time is not sufficient for another data transmission, it can not be transmitted within this TXOP but has to wait for the TXOP in the next SI. Let t_{DA_i} stand for the entire duration for a successful data transmission for RTSN$_i$, which is given by

$$t_{DA_i} = E[t_{DATA_i}] + t_{SIFS} + t_{ACK} + t_{SIFS} \tag{6}$$

If an MSDU arrives at the head of queue within the last t_{DA_i} of a TXOP, it is not allowed to get transmitted until experiencing another non-TXOP duration. Fig. 3 shows the relationship among TXOP, t_{DA_i} and non-TXOP. Under unsaturated and saturated conditions, if the MSDU arrives inside the first $t_{TXOP_i} - t_{DA_i}$ of a TXOP, it can get transmitted within the current TXOP. This is because when considering the unsaturated and saturated conditions, there is no MSDU buffered in the queue at the instant of $t_{TXOP_i} - t_{DA_i}$ of each TXOP. The reason is that transmission capability of the reserved bandwidth in these two cases is at least not less than actual traffic load of the corresponding RTSN. For simplify the equations hereafter, the duration of the first $t_{TXOP_i} - t_{DA_i}$ within a TXOP is represented by Φ_i. The first MSDU that arrives since the time of Φ_i will reach the head of the line and then it waits for the consequent dedicated bandwidth. In this case, the channel access delay is the duration between the moment that the MSDU reaches the head of the line and the beginning of the next TXOP. If an MSDU generates after the time when the first MSDU reaches the head of queue within the interval $[\Phi_i, \Phi_i + \mu_i]$ of the current SI, it will buffer in the queue and reach the head of the line until its prior MSDUs finish their transmissions in the subsequent reserved time-slots. As a result, these MSDUs have no channel access delay. The channel access delay for MSDUs of RTSN$_i$ in different conditions can be formulated by

$$dc_{j,i} = \begin{cases} 0, & if \;\; 0 \leq \eta(j) \leq \Phi_i \\ (1 - \frac{\eta(j)}{\Delta SI}) \cdot \Delta SI, & if \;\; \Phi_i < \eta(j) \leq \mu_i + \Phi_i \\ 0, & if \;\; \eta(j) > \mu_i + \Phi_i \end{cases} \tag{7}$$

To simplify the computation of average channel access delay for a QoS guaranteed RTSN, a period can be identified for $\eta(j)$. The proof is shown as follows.

Proof: Assumed that there is an integer P, which denotes the subsequent Pth MSDU that arrives after the jth MSDU. The offset duration of the Pth MSDU can be derived by

$$\begin{aligned}\eta(j+P) &= \delta(\frac{(j+P)\cdot\mu_i}{\Delta SI})\cdot\Delta SI \\ &= \delta(\frac{j\cdot\mu_i}{\Delta SI}+\frac{P\cdot\mu_i}{\Delta SI})\cdot\Delta SI\end{aligned} \tag{8}$$

The SI can be deemed as a fixed set of $SlotTime\ \sigma$, which is denoted by

$$\Delta SI = K\cdot\sigma, \quad if\ \ K\in\mathbf{N}^+ \tag{9}$$

Similarly, we can obtain

$$\mu_i = K'\cdot\sigma, \quad if\ \ K'\in\mathbf{N}^+ \tag{10}$$

where sending interval is expressed by an integer amount of σ. Using (9) and (10), the variance of (8) can be derived by

$$\eta(j+P) = \delta(\frac{j\cdot\mu_i}{\Delta SI}+\frac{P\cdot K'}{K})\cdot\Delta SI \tag{11}$$

Note that K and K' are both taken as integer values. Consequently, a minimum value of integer P can be found in order to make the value of $\frac{P\cdot K'}{K}$ equal to a positive integer. Due to the property of $\eta(x)$, the term $\frac{P\cdot K'}{K}$ can be ignored. Thus,

$$\eta(j) = \eta(j+P) \tag{12}$$

This verification suggests that the offset value of an arbitrary MSDU will periodically reappear after a certain duration which can be viewed as a period. Therefore, the average channel access delay can be obtained through computing the average value of channel access delay for all the MSDUs arrived within a period. The average channel access delay is expressed by

$$\overline{d_{ca,i}} = \frac{\sum_{j=1}^{p_i} dc_{j,i}}{p_i} \tag{13}$$

where p_i denotes the minimum period for the MSDUs of RTSN_i.

Queueing Delay. The queueing delay is measured from the moment that an MSDU pumps into the interface queue to the instant that it reaches the head of the line. Considering unsaturated and saturated conditions, the buffered MSDU during a SI can be completely transmitted using the TXOP in the subsequent SI. As a result, there is no MSDU buffered at the time of Φ_i in each TXOP. As shown in Fig. 6 and Fig. 7, there are three conditions of queueing delay for an MSDU.

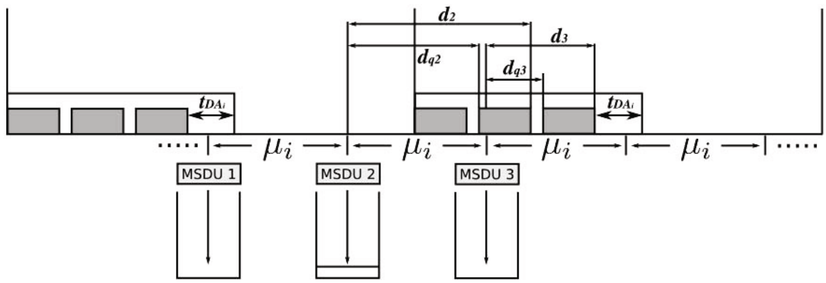

Fig. 6. Queueing delay on different time-slots

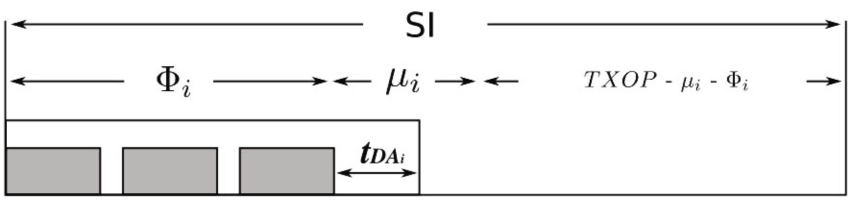

Fig. 7. Queueing delay on different conditions

First, the MSDU arrives inside the interval $[\Phi_i, \Phi_i + \mu_i]$ of the current SI. Second, the MSDU arrives within the interval $[\Phi_i + \mu_i, \Delta SI]$ of the current SI. Third, the MSDU arrives during the first Φ_i of the TXOP within the current SI.

In the case of the first situation, the newly arrived MSDU directly becomes the head-of-line MSDU and defers until the start time of its next TXOP. Thus, the MSDU has no queueing delay. For the MSDU following the second situation, it will buffer in the queue and get transmitted in the subsequent TXOP. Therefore, its queueing delay is equal to the deferral time of non-TXOP duration plus the transmission time of the prior MSDUs buffered in the queue. The amount of MSDUs that arrive prior to the tagged MSDU is $\lfloor \frac{\eta(j)-\Phi_i}{\mu_i} \rfloor$. They will cost the transmission time of $\lfloor \frac{\eta(j)-\Phi_i}{\mu_i} \rfloor \cdot t_{DA_i}$, which is part of the queueing delay of the tagged MSDU. The rest part of its queueing delay is the non-TXOP duration which is $\Delta SI - \eta(j)$. For the third situation, the MSDU will buffer in the queue and get transmitted in the current TXOP. As a result, its queueing delay is the transmission time of the remained MSDUs buffered before plus the residual transmission time of the MSDU which is being transmitted at the moment. In order to figure out the queueing delay in this situation, the amount of accumulated MSDUs from the instant Φ_i of last TXOP to the arrival time $\eta(j)$ of the tagged MSDU needs to be figured out. Since there is no MSDU at the timestamp Φ_i of the last TXOP, the number of MSDUs that still buffer in the queue is $\lfloor \frac{\eta(j)+\Delta SI-\Phi_i}{\mu_i} \rfloor - \lfloor \frac{\eta(j)}{t_{DA_i}} \rfloor - 1$. It excludes the current transmitting MSDU, which requires the time of $\lfloor \frac{\eta(j)}{t_{DA_i}} + 1 \rfloor \cdot t_{DA_i} - \eta(j)$ in order to finish its transmission. The amount of MSDUs that have already been sent is $\lfloor\frac{\eta(j)}{t_{DA_i}}\rfloor$. Finally, the queueing delay for the third situation can be derived as $(\frac{\eta(j)+\Delta SI-\Phi_i}{\mu_i}\rfloor - \lfloor\frac{\eta(j)}{t_{DA_i}}\rfloor - 1)\cdot t_{DA_i} + (\lfloor\frac{\eta(j)}{t_{DA_i}}+1\rfloor\cdot t_{DA_i} - \eta(j))$. The queueing delay for all the situations can be denoted by (14). Using the periodicity property of

$$dq_{j,i} = \begin{cases} (\lfloor\frac{\eta(j)+\Delta SI-\Phi_i}{\mu_i}\rfloor - \lfloor\frac{\eta(j)}{t_{DA_i}}\rfloor - 1)\cdot t_{DA_i} + (\lfloor\frac{\eta(j)}{t_{DA_i}}+1\rfloor\cdot t_{DA_i} - \eta(j)), & if \quad \eta(j) < \Phi_i \\ 0, & if \quad \Phi_i \leq \eta(j) \leq \mu_i + \Phi_i \\ \lfloor\frac{\eta(j)-\Phi_i}{\mu_i}\rfloor\cdot t_{DA_i} + \Delta SI - \eta(j), & if \quad \eta(j) > \mu_i + \Phi_i \end{cases} \tag{14}$$

the $\eta(j)$, the average queueing delay of RTSN_i is expressed by (15).

$$\overline{d_{q,i}} = \frac{\sum_{j=1}^{p_i} dq_{j,i}}{p_i} \tag{15}$$

Transmission Delay. Transmission delay is equal to the duration from the instant that an MSDU begins accessing the channel to the moment it is successfully transmitted. The average transmission delay can be denoted by

$$\overline{d_{tr,i}} = E[t_{DATA_i}] + t_{SIFS} + t_{ACK} + t_{SIFS} \tag{16}$$

where $E[t_{DATA_i}]$ stands for the average transmission time of an MSDU for RTSN_i.

4 Optimization Study Based on Delay Bound for SF-Based RR Scheme

In this section, we study the optimization of system parameters such as SI and the size of allocated TXOP for each RTSN. It can be implied from the previous analysis that guaranteed QoS of a RTSN can be achieved under unsaturated and saturated conditions in which required scheduling rate of the RTSN does not exceed the maximum transmission capability of its reserved TXOP. According to (1), the size of TXOP allocated for a RTSN lies with the amount of MSDUs that are allowed to be transmitted. Considering the delay bound, the duration of TXOP for a RTSN is closely associated with the size of SI, required scheduling rate and the average size of an MSDU. RTSNs with different required scheduling rates need to obtain distinct amount of resources (i.e. TXOPs) in order to ensure their guaranteed QoS. The delay of a RTSN with large required scheduling rate may be bound through allocating sufficient bandwidth. However, it will make less RTSNs reserve TXOPs in CFP. The trade-off between the optimal amount of RTSNs accommodated in one SI and the guaranteed delay for these RTSNs is an open issue.

On the other hand, another trade-off exists between a small and a large SI. Small SI can enhance the maximum transmission capability of each allocated

TXOP so that it is capable of accommodating RTSNs with a higher required scheduling rate. Small SI also reduces the delay for each admitted RTSN because of its short non-TXOP duration. However, the amount of RTSNs that can reserve the bandwidth in CFP decreases if small SI is employed. Using a large SI, more resources can be allocated to CFP. But each RTSN may require more bandwidth for satisfying their delay bound due to the degraded performance caused by large SI. To balance this trade-off, the optimization study is a necessity.

The aim of the optimization study is to accommodate maximum amount of RTSNs in an optimum SI given that the QoS of each RTSN is guaranteed. To investigate the optimization, the requirement of the reserved bandwidth for guaranteeing QoS toward RTSNs needs to be identified first. In fact, each dedicated TXOP has its own maximum transmission capability. In order to formulate the maximum transmission capability of a reserved TXOP, the reserved scheduling rate $\lambda_{r,i}$ for RTSN$_i$ is introduced. It can be given by

$$\lambda_{r,i} = \frac{n_{t,i} \times \overline{s_{DATA_i}}}{\Delta SI} \tag{17}$$

The reserved scheduling rate stands for the maximum transmission capability of the allocated TXOP. The prerequisite of the optimization is that each RTSN need to obtain satisfactory QoS which can be embodied by the requirement of delay bound. Thus, the QoS demand can be defined by

$$d_{ave,i} \leq d_{rmax,i} \tag{18}$$

where $d_{rmax,i}$ stands for the delay bound for RTSN$_i$. Using the definition of delay in section 3.1, the variance of (18) can be expressed by

$$\overline{d_{ca,i}} + \overline{d_{tr,i}} + \overline{d_{q,i}} \leq d_{rmax,i} \tag{19}$$

It has been indicated that the guaranteed QoS can only be achieved under the unsaturated and saturated conditions. This argument can be converted to the relationship between the reserved scheduling rate and the required scheduling rate, which is denoted by

$$\lambda_{r,i} \geq \lambda_i \tag{20}$$

The above argument suggests that QoS demand of a RTSN can only be satisfied when the reserved scheduling rate exceeds the required scheduling rate. In order to find out the connection between the required scheduling rate and the size of TXOP which is denoted by TXOPlimit, the relationship between the reserved scheduling rate and the TXOPlimit needs to be investigated. The size of TXOP can be expressed by

$$t_{TXOP_i} = \frac{\overline{s_{DATA_i}} \cdot n_{t,i} + O_{t,i}}{R} \tag{21}$$

which indicates that an entire duration of a TXOP for RTSN$_i$ is consumed by transmissions of MSDUs, ACK, MAC header and deferral time SIFS. The transmissions of ACK, MAC header and the deferral time SIFS are deemed as the overhead denoted by $O_{t,i}$. It is expressed by

$$O_{t,i} = n_{t,i}(s_{ACK} + O_{mac}) + 2n_{t,i} \cdot t_{SIFS} \cdot R \tag{22}$$

where O_{mac} represents the MAC header. Using (17), the derivative of (21) is obtained by

$$t_{TXOP_i} = \frac{\lambda_{r,i} \cdot \Delta SI + O_{t,i}}{R} \tag{23}$$

Considering the requirement defined by (20), (23) can be transformed into

$$t_{TXOP_i} \geq \frac{\lambda_i \cdot \Delta SI + O_{t,i}}{R} \tag{24}$$

which stands for the relationship between the TXOPlimit and required scheduling rate under saturated and unsaturated conditions. Since multiple RTSNs can reserve bandwidth in CFP, the average TXOP $\overline{t_{TXOP}}$ is introduced and given by

$$\overline{t_{TXOP}} = \frac{1}{N}\sum_{i=1}^{N} t_{TXOP_i} \tag{25}$$

Provided that the size of CFP as well as the SI are fixed, the optimum network capacity can be obtained if the existing reserved TXOPs occupy the minimum duration in CFP given that the delay bound of each RTSN is strictly satisfied. Converting this argument to the average TXOP, the network capacity can be maximized if the $\overline{t_{TXOP}}$ gets the minimum value. The expression is derived by

$$\begin{aligned} f(\lambda_{r,i}) &= min\{\overline{t_{TXOP}}\}, \;\; if \;\; d_{ave,i} \leq d_{rmax,i} \\ &= min\{\frac{\sum_{i=1}^{N} t_{TXOP_i}}{N}\}, \;\; if \;\; d_{ave,i} \leq d_{rmax,i} \end{aligned} \tag{26}$$

Using (23), the equation of optimum resource allocation can be finalized by

$$f(\lambda_{r,i}) = min\{\frac{\sum_{i=1}^{N}(\lambda_{r,i} \cdot \Delta SI + O_{t,i})}{N \cdot R}\}, if \; d_{ave,i} \leq d_{rmax,i} \tag{27}$$

where the reserved scheduling rate is the variable. It can be indicated from (27) that under the delay bound and fixed SI, the optimum resource allocation as well as the optimum network capacity can be achieved once the reserved scheduling rate of each RTSN is taken as the minimum value. It can be derived from (24) that if delay bound is sufficed, the lowest value of (27) is $\frac{\sum_{i=1}^{N}(\lambda_i \cdot \Delta SI + O_{t,i})}{N \cdot R}$ when $\lambda_{r,i}$ is equal to λ_i.

After figuring out the optimum resource allocation, the optimization of the SI can be investigated. Given the amount of QoS guaranteed RTSNs, if the size of SI changes, the reserved TXOPs of these RTSNs will vary accordingly in order to satisfy distinct level of delay bound. The optimum SI is obtained when the proportion of the reserved TXOPs for these RTSNs in CFP takes the minimum value. Therefore, the optimum SI can be formulated by

$$g(SI) = min\{\frac{\sum_{i=1}^{N} t_{TXOP_i}}{\Delta SI}\}, \quad if \; d_{ave,i} \leq d_{rmax,i} \tag{28}$$

where SI is the variable of the optimization. When the optimal SI is achieved, maximum duration can be left in SI for accommodating more RTSNs. Thus, maximizing the network capacity.

5 Performance Evaluation

In this section, we apply the analytical model to predict the delay performance of RTSNs in two SF-based RR schemes which are EDCA/RR and EDCA/DRR. Simulation and analytical outcomes are compared in order to verify the accuracy of the proposed mathematic model. Optimization results are also shown and discussed. Tab. 1 recaps all the parameters used in the evaluation. Several senders with a RTSN per-node as well as a receiver comprise the network. All the nodes are randomly deployed within each other's transmission range.

Table 1. Simulation parameters

Parameter(units)	Value
SIFS(μs)	10
Slot time(μs)	20
ACK size(bytes)	28
MAC header(bytes)	36
Channel capacity(Mbps)	11
Interface queue size(packets)	50
Transmission range (m)	250
Traffic application	CBR over UDP

Fig. 8 shows the simulation and analytical results of delay performance for QoS guaranteed RTSNs in EDCA/RR. The analytical results well-match the simulation outcomes. In order to study the impact of SI as well as the required scheduling rate on delay, three distinct size of SIs (i.e. 10ms, 15ms, and 20ms) are used and the required scheduling rate varies from 300kb/s to 1500kb/s. In this case, the allocated resource is made equal to the required bandwidth. It can be concluded that given a SI, delay of RTSNs with TXOPs become higher along

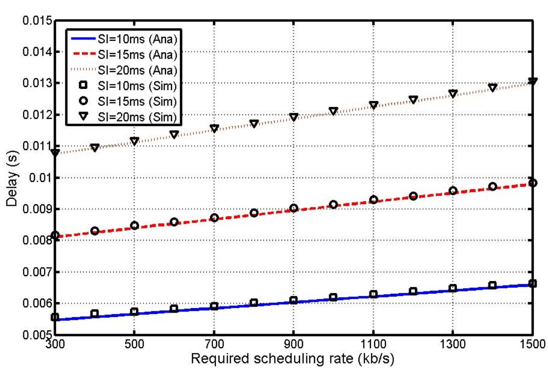

Fig. 8. Delay of RTSNs in EDCA/RR

with the increment of the required scheduling rate. The reason is that average queueing delay increases once the number of MSDUs allowed to be transmitted in a TXOP becomes larger. It also can be seen from Fig. 8 that compared with smaller SI, delay leaps dramatically in the face of larger size of SI. This is because the increasing channel access delay contributed by the non-TXOP duration plays the major impact on the average delay.

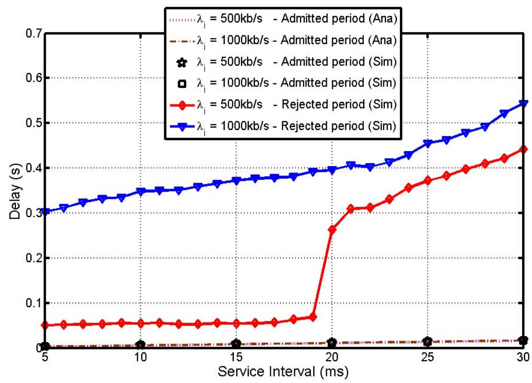

Fig. 9. Delay of RTSNs in EDCA/DRR

Fig. 9 shows the delay performance of RTSNs in EDCA/DRR. Two different required scheduling rates (i.e. 500kb/s and 1000kb/s) are used and the delay of RTSNs is tested under different size of SIs. Sufficient bandwidth is assigned to CFP (i.e. maximum CFP duration is equal to 0.75SI) in order to accommodate more RTSNs. On the other hand, a certain proportion of duration is allocated to CAP for the fairness toward other types of sessions or the rejected RTSNs which can not obtain dedicated resources in CFP. Simulation result implies that the rejected RTSN suffers from a degrade performance. This attributes to the channel contention from other NRTSNs and the limited duration for contending the chance of channel access. On the other hand, in the contention-based environment, rejected RTSN with higher required scheduling rate receives a worse delay performance than the rejected RTSN with lower required scheduling rate. This is because the buffered MSDUs in the interface queue accumulates faster in terms of the RTSN with higher required scheduling rate. In contrast to their performance in CAP, the delay drastically decreases if the RTSNs are re-admitted by the admission control algorithm in EDCA/DRR and transmitted during CFP when certain bandwidth in CFP becomes idle. In addition, a good agreement between the analytical outcomes and simulation results can be seen in Fig. 9.

As analysed in the previous section, the reserved scheduling rate of a RTSN needs to be at least equal to the required scheduling rate so that its QoS demand such as the delay bound can be possibly satisfied. Even if the above requirement is met, a RTSN may need a higher reserved scheduling rate compared with the

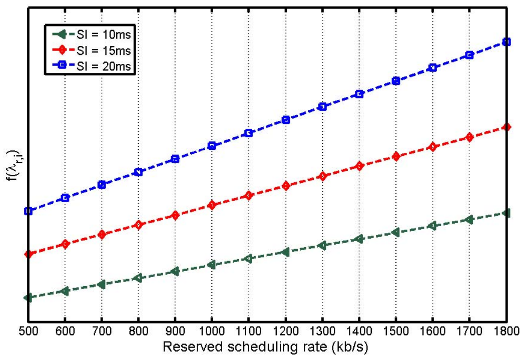

Fig. 10. Optimum scheduling rate

required scheduling rate if it is desired by a lower delay bound. Delay reduces if more bandwidth can be reserved by the corresponding RTSN. Consequently, to consider the optimization of system parameters such as the allocated resource (i.e. TXOP) and SI, the bandwidth reserved for each RTSN need to be the minimum value that can exactly meet the delay bound under certain value of SI. Fig. 10 shows the optimum scheduling rate of RTSNs with required scheduling rate of 500kb/s. It suggests that under a relatively loose delay bound, the optimum scheduling rate is equal to the required scheduling rate. From Fig. 10, we can also conclude that the more resources allocated to these RTSNs, the worse the network capacity becomes. However, reserved scheduling rate needs to be adjusted according to the QoS requirement (i.e. delay bound and required scheduling rate). To investigate the optimum SI under different QoS demand, a set of required scheduling rates (i.e. 200kb/s, 800kb/s, 1400kb/s and 2000kb/s) are given under the delay bound of 25ms. The SI is tested from 5ms to 100ms

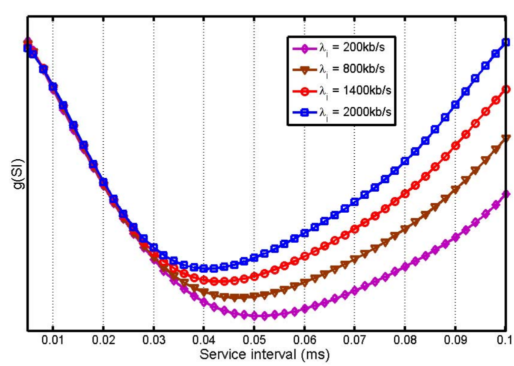

Fig. 11. Optimum SI

in order to identify the optimum value. Fig. 11 shows the results that have been figured out for all the different conditions. It can be seen that the optimum SI tends to be larger if RTSNs with lower required scheduling rates are employed. The reason is that under a light traffic load, RTSNs can reserve a relatively small size of TXOP which can satisfy their QoS. As the SI is increasing, the requested bandwidth will not leap dramatically so that the optimum value can be achieved at a larger value. The results shown in Fig. 10 and Fig. 11 validate the feasibility of the proposed optimization study.

6 Conclusion

Resource reservation plays an important role in QoS provisioning for multimedia applications in IEEE 802.11-based wireless networks. This paper provides an analytical model and an optimization study for SF-based distributed RR mechanisms. The proposed model has been proved accurate for evaluating the delay performance of RTSNs with TXOPs under different traffic conditions. In addition, optimization of system parameters such as the size of allocated TXOPs and the SI has been conducted for maximizing the network capacity, making more RTSNs obtain dedicated bandwidth in CFP. Simulation and analytical results have verified the accuracy of the analytical model. In addition, the analytical results for optimization have shown optimum parameters under different situations, which has proved the feasibility of the optimization study.

Acknowledgement. This work was performed in the project QoSMOS which has received research funding from the EU FP7 framework.

References

1. IEEE 802.11-2007: IEEE Standard for Wireless LAN Medium Acess Control (MAC) and Physical Layer (PHY) Specification (June 2007) (revised version)
2. Ghazizadeh, R., Fan, P.: Dynamic Priority Scheduling Mechanism through Adaptive InterFrame Space. In: Proc. IEEE WiCOM 2007, Shanghai, China (September 2007)
3. Naoum-Sawaya, J., Ghaddar, B., Khawam, S., Safa, H., Artail, H., Dawy, Z.: Adaptive Approach for QoS Support in IEEE 802.11e Wireless LAN. In: Proc. IEEE WiMOB 2005, Montreal, Canada (August 2005)
4. Kim, M.S., Shrestha, D.M., Ko, Y.B.: EDCA-TM: IEEE 802.11e MAC Enhancement for Wireless Multi-Hop Networks. In: Proc. IEEE WCNC 2009, Budapest, Hungary (April 2009)
5. Feng, Z., Wen, G., Zou, Z., Gao, F.: RED-TXOP Scheme for Video Transmission in IEEE 802.11e EDCA WLAN. In: Proc. IEEE ICCTA 2009, Beijing, China (October 2009)
6. Zhang, Y., Foh, C.H., Cai, J.: An On-Off Queue Control Mechanism for Scalable Video Streaming over the IEEE 802.11e WLAN. In: Proc. ICC 2008, Beijing, China (May 2008)

7. Hamidian, A., Körner, U.: Extending EDCA with Distributed Resource Reservation for QoS Guarantees. J. Sel. Telecommun Syst. 39(3-4), 187–194 (2008)
8. Yu, X., Navaratnam, P., Moessner, K.: Distributed Resource Reservation Mechanism for IEEE 802.11e-Based Networks. In: Proc. IEEE VTC 2010, Ottawa, Canada (September 2010)
9. Lin, Y., Wong, V.W.S.: Frame Aggregation and Optimal Frame Size Adaptation for IEEE 802.11n WLANs. In: Proc. IEEE GLOBECOM 2006, San Francisco, USA (December 2006)