



Article

Smart Preliminary Channel Access to Support Real-Time Traffic in Wi-Fi Networks

Kirill Chemrov , Dmitry Bankov * , Evgeny Khorov and Andrey Lyakhov

Institute for Information Transmission Problems of the Russian Academy of Sciences, 127051 Moscow, Russia

* Correspondence: bankov@iitp.ru

Abstract: Real-time applications (RTA) are an important use case for IEEE 802.11be, a new amendment to the Wi-Fi standard. This amendment introduces new complicated mechanisms to provide low delay and high reliability for RTA, but many of them are not supported by legacy devices that may be present in future Wi-Fi networks. In contrast, the preliminary channel access (PCA) method is designed to satisfy strict RTA requirements even in the presence of legacy devices and does not require significant changes to the Wi-Fi protocol. However, it significantly reduces the capacity for non-RTA traffic. This paper introduces a Smart PCA method, which improves the performance of all the stations in scenarios with multiple RTA stations. Extensive simulation shows that the Smart PCA method guarantees low delays for intensive RTA traffic in these scenarios. Moreover, it doubles the network capacity for the stations with non-RTA traffic.

Keywords: real-time applications; Wi-Fi; 802.11; low delay; reliability; preliminary channel access



Citation: Chemrov, K.; Bankov, D.; Khorov, E.; Lyakhov A. Smart Preliminary Channel Access to Support Real-Time Traffic in Wi-Fi Networks. *Future Internet* **2022**, *14*, 296. <https://doi.org/10.3390/fi14100296>

Academic Editor: Sachin Sharma

Received: 9 September 2022

Accepted: 14 October 2022

Published: 16 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The popularity of cloud gaming, virtual and augmented reality, video streaming services, and other real-time applications (RTA) is rapidly growing. These applications impose strict requirements on delay and reliability [1], which make RTA support in wireless networks especially challenging, though much effort has been expended in this direction.

In cellular systems, this problem is known as providing Ultra-Reliable Low Latency Communications [2–6], which is, perhaps, the most revolutionary part of 5G systems. To address this problem in Wi-Fi networks, the IEEE 802.11 Working Group, which develops Wi-Fi standards, has created an RTA Topic Interest Group (RTA TIG). RTA TIG has classified RTA scenarios and defined the quality of service (QoS) requirements for such traffic [7] as the delivery of 99.9–99.999% of packets with a delay below 1–10 ms. Providing such low delays and high reliability is extremely challenging in Wi-Fi networks because of the unlicensed spectrum and Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). During a long transmission, none of the stations (STAs) in the neighborhood can access the channel. Moreover, as Wi-Fi does not support full-duplex operation, a STA that has started a long non-RTA transmission (hereinafter, a non-RTA STA) cannot be interrupted by another STA willing to deliver an RTA frame (RTA STA). In addition, random channel access causes frame collisions, which degrade transmission reliability.

RTA traffic can be categorized into predictable and unpredictable traffic, and they require different solutions. Solutions for unpredictable traffic are proposed in many works [8–13]. In this paper, we focus on predictable traffic and attempt to exploit its features to serve RTA most efficiently. Such traffic can be found in many scenarios, such as virtual reality, industrial automation, etc.

Despite the difficulties described above, the support for RTA has become one of the targets of the future 802.11be amendment to the Wi-Fi standard, also known as Wi-Fi 7 [14]. The methods enabling RTA proposed in 802.11be use scheduling and coordination by the access point (AP) of how the STAs access the channel. However, legacy STAs do not

support these methods, which may make them useless in heterogeneous networks with STAs of various generations. Moreover, some methods, such as restricted Target Wake Time (r-TWT), are optional, and the STAs, even the 802.11be ones, that do not support them may violate the prioritization of RTA traffic. To support RTA in such networks, backward compatible and universal methods will be developed and investigated.

In this paper, we investigate such methods based on changing the channel access parameters and preliminary channel access (PCA) [15]. The PCA method considers accessing the channel in advance before RTA data appearance to reserve the channel for RTA transmissions. However, in scenarios with many RTA STAs, the total duration of reservations becomes rather long, and it significantly degrades the channel usage *efficiency* by non-RTA STAs, which we define as the portion of time when non-RTA STAs successfully transmit their payload. The main contribution of the paper is the enhanced method of PCA, called Smart PCA. In scenarios with many RTA STAs and given delay quantile limit for RTA STAs, Smart PCA allows increasing the efficiency of non-RTA STAs. In contrast with optional standard mechanisms for RTA prioritization, Smart PCA uses the basic NAV mechanism, mandatory for Wi-Fi devices since the very first generation. Those STAs that do not support Smart PCA cannot prevent RTA STAs from delivering their packets in time but obtain lower throughput. With simulation, we compare the reliability and delays of RTA frame transmission with various methods, as well as the efficiency of non-RTA STAs. As a result, we determine the parameters of RTA traffic and its QoS requirements for which it is advisable to use the proposed method.

The rest of the paper is organized as follows. Section 2 contains a brief description of the basic channel access method in Wi-Fi networks. In Section 3, we describe RTA-aware solutions in Wi-Fi networks known in the literature. Section 4 introduces the investigated methods of RTA traffic service, including the Smart PCA method proposed in this paper. Section 5 discusses the obtained numerical results. Section 6 concludes the paper.

2. Channel Access in Wi-Fi Networks

Modern Wi-Fi networks use Enhanced Distributed Channel Access (EDCA), which works as follows. EDCA defines four access categories (ACs): Voice, Video, Best Effort, and Background, each of which corresponds to a separate queue of packets at every STA. If a new frame arrives in an empty queue while the channel is busy, the STA initializes the backoff counter with a random integer uniformly distributed in $[0, W_{min} - 1]$, where W_{min} is the minimum size of the contention window.

If the channel is busy, the STA suspends its backoff counter. When the channel becomes idle and remains idle for the Arbitration Interframe Space (AIFS) time, the STA resumes and decrements its backoff counter by one. Then, the STA decrements it by one each time the channel is idle for the time T_e . The STA transmits the frame if the counter is equal to zero, and a new decrement time comes. If a new frame arrives in an empty queue when the channel has been idle for the AIFS time and the backoff counter after the previous transmission has become zero, then the STA immediately transmits the frame. In case of successful reception, the receiving STA responds with an acknowledgment (ACK) frame after the Short Interframe Space (SIFS) time. If a transmitting STA does not receive an ACK within the ACKTimeout interval since the end of its transmission, it considers the transmission unsuccessful and repeats channel access with a new backoff initialized with a random value selected from a larger interval.

Specifically, after the failed transmission attempt r , the contention window doubles until it reaches the maximum value W_{max} :

$$W_r = \begin{cases} W_{min}, & r = 0; \\ \min\{2W_{r-1}, W_{max}\}, & 0 < r < R. \end{cases}$$

The frame is discarded if the number of failed attempts r reaches the retry limit R .

For each AC, the STA uses different channel access parameters. Table 1 shows the W_{min} , W_{max} , AIFSN, and TXOP limit, where AIFSN defines the duration of the AIFS = SIFS + $T_e \cdot$ AIFSN interval and can be an integer from 2 to 15, and the TXOP Limit sets the maximum time the STAs can occupy the channel. The standard defines the default values of these parameters, but the AP can change them.

Table 1. Default EDCA parameters for client STAs.

AC	Wmin	Wmax	AIFSN	TXOP Limit
Voice	4	8	2	2.08 ms
Best Effort	16	1024	3	2.528 ms

The STAs also use the Request-to-Send/Clear-to-Send (RTS/CTS) mechanism, which works as follows. Before data transmission, the STA sends an RTS frame. Then, if the receiving STA considers the channel idle, it responds with a CTS frame. Only after receiving CTS, does the STA begin data transmission. If the STA does not receive the CTS frame within the CTSTimeout interval, it retransmits RTS. Other STAs that have received at least one of these frames consider the channel virtually busy until the end of the data transmission and acknowledgment, which is implemented as follows. Both RTS and CTS contain the Duration/ID field that specifies for how long the channel is busy after the frame is transmitted. STAs that read this field set the Network Allocation Vector (NAV) counter according to the specified time and consider the channel busy until the counter is reset to zero. Although the RTS/CTS handshake adds overhead, collisions between two RTS frames are much shorter than collisions between long data frames, so the RTS/CTS mechanism can improve delays and throughput.

3. Related Works

Solutions for RTA traffic service in Wi-Fi networks are studied in the IEEE 802.11 working group and in the literature.

Some RTA-aware solutions rely on mechanisms that appeared in the 802.11ax standard amendment [16,17], e.g., Orthogonal Frequency Division Multiple Access (OFDMA) [18–22]. OFDMA allows the AP to allocate different subcarriers of the same channel to multiple STAs for simultaneous reception or transmission of data. An OFDMA-based approach to RTA traffic service was explored in [8] and presented in RTA TIG. Note that to allocate frequency resources in OFDMA, the AP needs to know when and which STAs have data. As a result, the transmission of information about the appearance of frames in the queue causes additional delays. To avoid this problem, the AP can leave a part of the channel for random access [23–29] so that an STA can attempt to transmit as soon as it has a frame. However, with random access, STAs transmit data according to the access method that is similar to the ALOHA [30,31], the efficiency of which significantly reduces with a large number of contending STAs.

The 802.11ax standard also defines the Quiet Time Period (QTP) mechanism [32], which allows STAs to request time intervals during which other STAs are not allowed to access the channel. QTP is important for RTA because transmissions in such time intervals avoid contention-related delays. For example, the authors of [33] propose an RTA traffic service approach based on periodic reservations implemented with QTP. With this approach, the AP periodically allocates time intervals for transmissions of only one of the RTA STAs. However, the QTP mechanism is optional, and the presence of STAs not supporting it may violate the restrictions on low delays. In addition, if several Wi-Fi networks overlap, this approach requires an additional central coordinator, which distributes time among the networks.

As RTA traffic service is an important use case for Wi-Fi 7 networks, many works are devoted to the study of the mechanisms introduced in the 802.11be amendment. Overviews of Wi-Fi 7 innovations are included in [34–38], and their application for serving delay-

sensitive traffic is discussed in [39]. The standard defines the restricted target wake time (r-TWT) mechanism to reserve channel resources and provide low delays for predictable and delay-sensitive traffic. However, the mechanism becomes unreliable in the presence of legacy devices that do not support r-TWT and have poor clock accuracy. Moreover, r-TWT is an optional mechanism, and the standard states [14] (Section 35.9.4.2) that the Wi-Fi 7 STAs that do not support this method can ignore the r-TWT service periods, which violates the prioritization of RTA packets. This behavior is favorable for non-RTA STAs and implicitly makes the STAs not support this method if they do not have RTA traffic.

The 802.11be standard also introduces a multi-link operation (MLO) technology that can be used to serve RTA. Notably, some RTA-aware ideas based on simultaneous work in several channels had been studied even before MLO became a part of Wi-Fi. For instance, the paper [40] suggests using an additional radio interface to maintain a control channel. In this proposal, a STA sends a busy tone in the control channel when an RTA frame is enqueued. All STAs listen to the control channel and will immediately stop non-RTA transmissions once they detect the busy tone. However, low delays cannot be guaranteed in the presence of STAs that do not support such a mechanism. Another example of RTA traffic service using several channels is introduced in [41–43]. The authors study a method of reducing channel access delay in which STAs access several channels simultaneously and transmit in the first accessed one. The authors of [44] propose an adaptation of this method for MLO. When a multi-link device generates a packet, it duplicates it and contends for access in multiple channels. If the device receives an acknowledgment on at least one channel, it removes the copies of the packet from the queues of other channels. This approach reduces the channel access delay because the delay is determined by its minimal value among multiple channels. In [9–12,45,46], the authors also explore the potential of MLO to reduce delays. Note that all the methods studied in these works require the PHY complication of RTA devices, in particular, to have at least one extra interface for wireless medium access, which significantly increases the cost of devices.

Some works consider simple ways to reduce delays. For example, the authors of [13] introduce a new AC for delay-sensitive traffic. The parameters of this AC will guarantee that RTA STAs win the contention with non-RTA STAs. However, this method cannot reduce the delay associated with waiting for the end of non-RTA transmission. Other approaches consider predictable RTA traffic. In this case, RTA STAs expect the arrival of data and can start accessing the channel in advance. Thus, the authors of [47] propose to start the backoff countdown in advance before the RTA frame appears at the STA to reduce the channel access time. However, with this solution, the STA loses channel access if the data frame has not arrived by the moment when the backoff countdown is over. To address this drawback, in [15] we propose the preliminary channel access (PCA) method, according to which the STA reserves the channel by sending an RTS frame. As a result, the STA that successfully transmits an RTS frame but has no data frame does not lose channel access until an RTA data frame is delivered. We also consider reducing the TXOP Limit for delay-insensitive traffic as an alternative method to reduce the delay associated with waiting for the end of a non-RTA transmission. The paper shows that in the case of rare RTA traffic, the PCA method is more efficient than the reduction of the TXOP limit. However, the study is carried out in the scenario with only one RTA STA in the network. Increasing RTA STAs numbers can cause a significant increase in delay due to contention between STAs and waiting for the reservation time of another RTA STA.

To sum up, many of the existing methods for RTA traffic service in Wi-Fi networks either require STAs to have a more complex PHY or cannot guarantee reliability in the presence of legacy STAs. At the same time, approaches without these disadvantages are compared only in scenarios limited to one RTA STA in the network.

In this paper, we state a problem to develop an improvement of the PCA method, which makes it possible to increase the channel usage efficiency of non-RTA STAs providing a given delay quantile limit for RTA STAs. The improved method should not complicate the PHY of RTA STAs and should work if other STAs in the network do not support the new Wi-

Fi standard. We also state a problem to compare the improved method with the previously proposed ones and to determine the traffic parameters for which the improvement is more efficient to use.

4. Investigated Methods to Support RTA

Consider a Wi-Fi network consisting of an AP, RTA STAs, and non-RTA STAs. Let all non-AP STAs generate uplink traffic. In this case, we distinguish the following components of the RTA frame transmission delay: waiting for the channel to become idle, competing with RTA and non-RTA STAs, and transmission of the data frame, including ACK reception. If the RTA traffic intensity is low and the RTA transmissions are short, then the main contribution to the delay is made by contention with non-RTA STAs and waiting for the end of their transmissions. To minimize the delay component associated with contention with non-RTA STAs, we set the EDCA parameters for ACs corresponding to RTA STAs and non-RTA STAs as follows. The minimum and maximum size of the contention window will not exceed the difference between the smallest AIFSN for non-RTA traffic and AIFSN[RTA]:

$$W_{min}[RTA] \leq W_{max}[RTA] \leq \min\{AIFSN[non-RTA] - AIFSN[RTA]\}.$$

Such parameters guarantee that RTA STAs always win the contention with non-RTA ones.

Next, we consider three methods to reduce the delay component associated with waiting for the end of a non-RTA transmission.

4.1. Tuned EDCA

The first method is Tuned EDCA [15]. It involves reducing the TXOP Limit for the AC corresponding to non-RTA STAs. Note that reducing the time for which non-RTA STAs can occupy the channel decreases the payload portion in one transmission, and, as a result, decreases the throughput. Therefore, the TXOP Limit will not be reduced below the value that is enough to provide the required delays for RTA.

4.2. Preliminary Channel Access (PCA)

The second method is PCA [15], shown in Figure 1. Let the RTA STA expect the arrival of a data frame during some interval T_{exp} . According to PCA, the RTA STA starts channel access T_{PCA} before the beginning of this interval to transmit an RTS frame. Let us denote the duration of the RTA data frame transmission, including the acknowledgment, by T_s . In case of timely RTS transmission, the AP responds with CTS, and the channel becomes virtually busy for all other STAs until the end of the possible transmission, i.e., $T_{exp} + T_s$ interval. As a result, the STA can transmit a data frame immediately after its generation. Otherwise, if the data frame arrives before RTS transmission, the STA discards the RTS frame and starts the channel access to transmit the data frame. To avoid waste of channel resources, when the data frame transmission ends before the end of $T_{exp} + T_s$ interval, the STA transmits a CFend frame which signals all STAs to zero their NAVs.

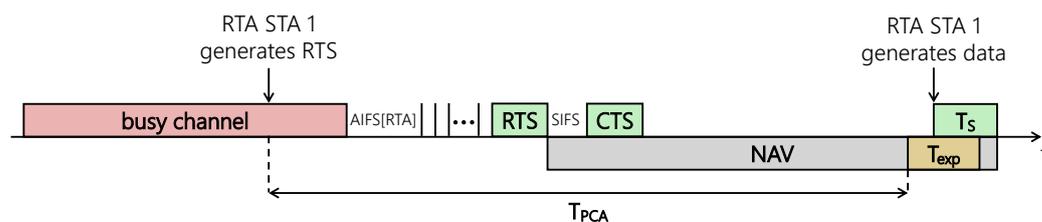


Figure 1. Scheme of the PCA.

The time T_{PCA} for which an RTA STA starts PCA equals

$$T_{PCA} = TXOP\ Limit + AIFS[RTA] + W_{min}[RTA] \cdot T_e + T_{RTS} + SIFS + T_{CTS}.$$

This T_{PCA} value guarantees a successful channel reservation by the data frame arrival time if only one RTA STA starts PCA (generates an RTS frame) while the channel is busy with a non-RTA transmission. The disadvantage of the PCA is the fact that one RTA STA can occupy the channel for a long time while other RTA STAs have data frames to transmit. As a result, delay components associated with waiting for the end of RTA transmission and contention with RTA STAs become significant.

4.3. Smart Preliminary Channel Access (Smart PCA)

This paper proposes an improved PCA which we call Smart PCA and which works as follows. If the number M of RTA STAs in the network is one or the maximum channel reservation time is less than the duration of one RTA transmission T_s :

$$\begin{cases} M = 1, \\ T_{PCA} - (T_{RTS} + SIFS + T_{CTS}) < T_s; \end{cases} \tag{1}$$

then RTA STAs access the channel in the same way as in the PCA. Otherwise, the modification of the PCA is as follows.

The RTA STA starts PCA by generating an RTS frame $T_{SmartPCA}$ before the beginning of T_{exp} interval. SIFS after receiving the CTS frame from the AP, the RTA STA transmits a special frame, which we call SPCA. It has almost the same structure as the RTS frame but has additional fields containing information about the beginning and the duration of the T_{exp} interval, see Figure 2. After receiving the SPCA frame, other RTA STAs with data can access the channel, but only if their transmission ends before the start of the T_{exp} interval. If the transmission fails, the RTA STAs can retry, but before each channel access, they will ensure that their transmissions end before the T_{exp} interval.

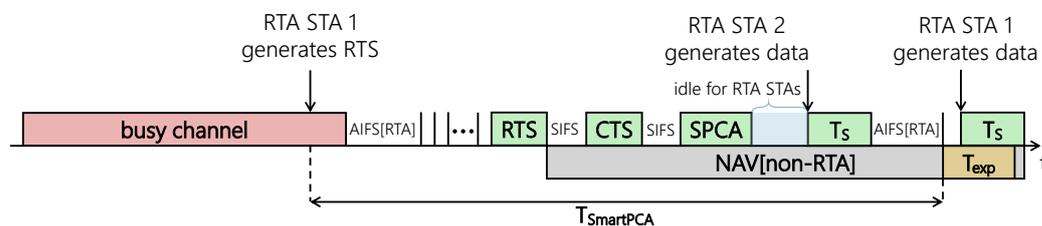


Figure 2. Transmission scheme of Smart PCA.

Compared to the PCA, in this method, the RTA STA starts channel access earlier due to the transmission of an additional SPCA frame:

$$T_{SmartPCA} = T_{PCA} + SIFS + T_{SPCA}.$$

The SPCA frame consumes additional channel time, and therefore it is useless if condition (1) holds because there are either no other RTA STAs, or they are unable to finish their transmission before T_{exp} interval. Thus, the Smart PCA guarantees that, in case of timely RTS transmission, RTA STA can transmit data frames as soon as they appear within the T_{exp} interval. At the same time, if condition (1) is not met, the improved method allows RTA STAs to reduce delays by transmitting data in the channel reserved by another RTA STA.

Note that with both PCA and Smart PCA, RTA transmissions, including the PCA procedure, consume more channel time than short RTA transmissions in the Tuned EDCA.

5. Results

To study the methods described in Section 4, we develop a discrete-event simulator similar to ns-3 [48]. We consider a Wi-Fi network with an AP, N non-RTA STAs, and M RTA STAs. All the STAs sense transmissions of each other perfectly, so there are no hidden STAs. Non-RTA STAs support the 802.11n standard, but without significant changes in

results, we can consider 802.11ac STAs. Non-RTA STAs operate in saturation mode and serve the traffic of one AC with parameters W_{min} , W_{max} , AIFSN and TXOP Limit. They always occupy the channel for the maximum time defined by TXOP Limit and use the RTS/CTS mechanism to increase throughput.

RTA STAs support the 802.11be standard and generate quasiperiodic traffic (Figure 3). The expected frame arrival time has a period T_{period} , and the real arrival time is normally distributed around the expected time with a standard deviation σ . This model of traffic can describe RTA that periodically generates data, but the real frame arrival time is random due to the implementation issues, such as clock inaccuracy or unpredictable operation system delays. Let us define the interval around the expected frame arrival time as $T_{exp} = 10\sigma$. This value of T_{exp} guarantees that more than 99.999% of data frames are generated inside T_{exp} .

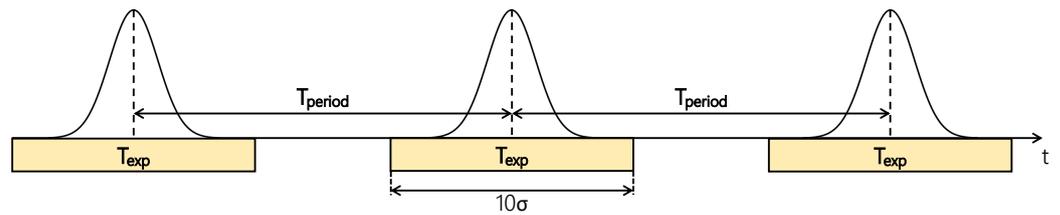


Figure 3. Quasiperiodic traffic of RTA STAs.

All the RTA STAs have the same periods T_{period} , but the cycles of different RTA STAs are shifted relative to each other by a random value uniformly distributed in $[0, T_{period}]$. The AC for RTA traffic has the following parameters: $W_{min}[RTA] = W_{max}[RTA] \triangleq W[RTA]$ and AIFSN[RTA]. We assume that all the RTA STAs use the same modulation and coding scheme, which is reliable enough to neglect frame losses because of random noise in the channel. The size of RTA MAC Protocol Data Units (MPDUs) is fixed and equal for all RTA STAs. Note that we consider a short duration of RTA frames; therefore, in the Tuned EDCA, RTA STAs do not use the RTS/CTS mechanism in order not to increase the delay because of the additional frames transmissions.

For the methods described in Section 4, we consider that all the STAs operate in a 40 MHz channel. For the multi-link method [44] described in Section 3, we assume that AP and RTA STA simultaneously operate in two 20 MHz channels, while N non-RTA STAs can operate only in one channel and are uniformly distributed among two 20 MHz channels. Note that as receiver sensitivity is higher in the narrower channel, RTA STAs in 20 MHz can use faster MCS than in 40 MHz without losing reliability.

The modeling parameters are listed in Table 2.

Table 2. Modeling parameters.

Parameter	Value	Parameter	Value
T_e	9 μ s	RTA MPDU size	2500 bytes
T_{CTS}	24 μ s	N	8
$SIFS$	16 μ s		
Parameter	Value for RTA STAs	Value for non-RTA STAs	
AIFSN	2	10	
W_{min}	8	16	
W_{max}	8	1024	
Standard	802.11be	802.11n	
OFDM Symbol	13.6 μ s	4 μ s	
Guard Interval	0.8 μ s	0.8 μ s	
Preamble	48 μ s	40 μ s	
MCS for 40 MHz	MCS 2	MCS 6	
MCS for 20 MHz	MCS 3	MCS 6	

To compare the RTA service methods, we consider the following QoS metrics. For RTA traffic, we use the *0.99999-quantile* of delay, which shows the minimum delay that can be guaranteed with a probability of 0.99999. The QoS metric for non-RTA traffic is the efficiency of channel usage by non-RTA STAs.

The further description is organized as follows. In Section 5.1, we consider how the described metrics depend on the number of RTA STAs. Here and further, we consider a fixed number N of non-RTA STAs. As shown in [15], the delay quantile for RTA STAs does not depend on N because for the considered methods EDCA parameters are such that the RTA STAs do not contend for the channel access with non-RTA ones, while the efficiency of channel usage for non-RTA STAs has the same decrease rate with the growth of N for Tuned EDCA and PCA, and the same result is valid for Smart PCA. Then, in Section 5.2 we show the influence of TXOP Limit on the metrics. Finally, in Section 5.3, we explore the tradeoff between the delay quantile and the efficiency of channel usage.

5.1. Impact of the RTA STAs Number

First, we consider the dependence of the metrics introduced above on the number M of RTA STAs. Figure 4 shows the simulation results for $T_{period} = 50$ ms, $\sigma = 10$ ms and TXOP Limit = 2 ms. With default EDCA, RTA STAs do not use any special method for RTA traffic service. In this case, RTA STAs and non-RTA STAs use EDCA parameters corresponding to AC Voice and AC Best Effort (Table 1). The obtained results show that with the default EDCA parameters, 0.99999-quantile of delay exceeds 20 ms, so RTA QoS requirements are violated. The multi-link method [44] described in Section 3 almost halves the delay quantile. However, RTA service methods described in Section 4 guarantee even lower delays and do not require devices to have multiple radio interfaces, so it only makes sense to compare these methods with each other. Note that the multi-link method can be combined with Tuned EDCA, PCA, or Smart PCA methods to further reduce the delay quantile, and it can be a direction of future research.

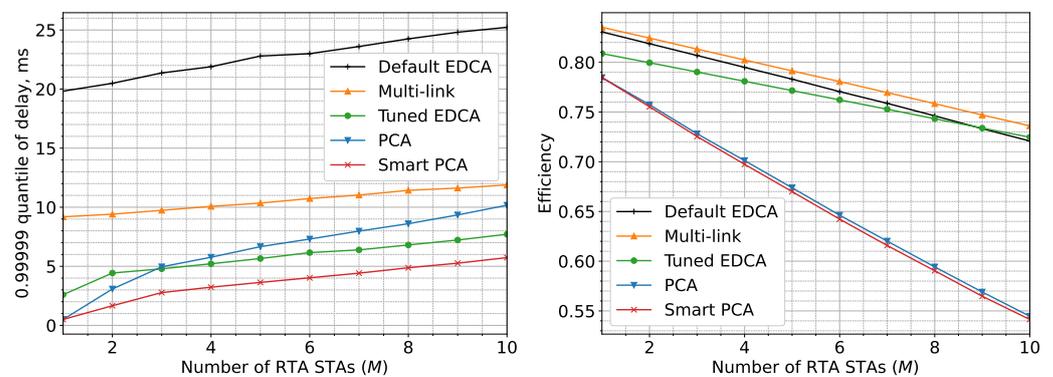


Figure 4. Delay quantile and efficiency vs. the RTA STAs number M in case of $T_{period} = 50$ ms, $\sigma = 10$ μ s, TXOP Limit = 2 ms.

For $M = 1$, the delay quantile for PCA and Smart PCA does not differ because condition (1) is met, so the channel access in both methods is the same. With increasing M , PCA loses not only to the Smart PCA but also to the Tuned EDCA. The reason is that in PCA, one RTA STA can occupy the channel for time $T_{PCA} + T_s$, which causes large delays for another RTA STA that generate data during this channel occupancy. In addition, before a successful transmission, an STA can lose contention to several other RTA STAs, which leads to waiting for several $T_{PCA} + T_s$ intervals. At the same time, when using Tuned EDCA, the main contribution to the delay is waiting only for TXOP Limit and maybe a few short T_s transmissions in case of losing to several RTA STAs. The Smart PCA retains the delay advantage of the Tuned EDCA with increasing M because instead of waiting for long channel occupancy by other RTA STAs, the RTA STA can transmit its data frame as soon as it arrives.

5.2. Impact of the TXOP Limit Value

Let us fix the number of RTA STAs: $M = 5$; the RTA traffic period: $T_{period} = 20$ ms; and $\sigma = 10$ μ s. Consider the dependences of delay quantile and efficiency on the TXOP Limit value. Figure 5 shows the monotonic increase in delay quantile for PCA and Tuned EDCA over the entire range of TXOP Limit. At the same time, for the Smart PCA, the quantile reaches a constant value starting from some value of TXOP Limit. With a long TXOP Limit, the reservation intervals are large enough, and other RTA STAs transmit data instead of waiting for them. In this case, condition (1) is not met, and the Smart PCA performs a constant delay. At low TXOP Limit, condition (1) is met, so the RTA STA does not have time to transmit in the channel reserved by another RTA STA. So, the growth of TXOP Limit causes an increase in the waiting time for another RTA STA reservation and hence an increase in delay.

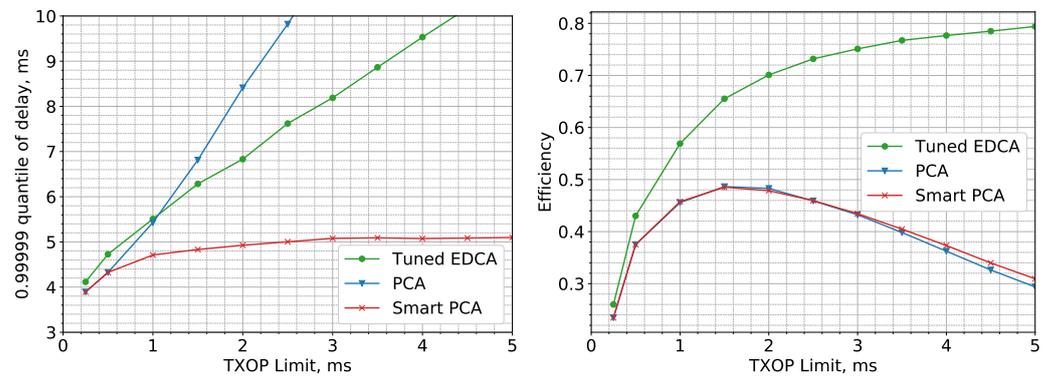


Figure 5. Delay quantile and efficiency vs. TXOP Limit in case of $T_{period} = 20$ ms, $\sigma = 10$ μ s, $M = 5$.

The efficiency of the Tuned EDCA increases monotonically because as the TXOP Limit increases, both the portion of successful transmissions and the portion of payload in one transmission increase. The efficiency in the PCA and the Smart PCA increases, reaches the peak, and then decreases, because in addition to the effect described above, with an increase in TXOP Limit, the values of T_{PCA} and $T_{SmartPCA}$ increase. As a result, RTA transmissions, including PCA procedure, take more channel time. A slight difference between the efficiency for the PCA and the Smart PCA is related to the difference in the values of T_{PCA} and $T_{SmartPCA}$.

5.3. Dependences of the Efficiency on the Delay Quantile

Let us compare the methods depending on the delay requirements. To do it, we consider how the efficiency of the studied methods depends on the delay quantile. Figures 6 and 7 show the results of six experiments for different values of RTA STAs number, RTA traffic period, and σ . To satisfy the requirement for the delay quantile (abscissa axis), it is necessary to set some TXOP Limit value, which in turn affects the efficiency of non-RTA STAs (ordinate axis). Thus, the plots in Figures 6 and 7, including the horizontal parts, show the maximum efficiency that each method provides for a given delay quantile limit.

First, we consider the strict requirements for the delay (3–5 ms). Figure 6a shows the simulation results for $T_{period} = 20$ ms, $\sigma = 10$ μ s, and $M = 5$. The maximum efficiency of the PCA and the Smart PCA is 48%. At the same time, Tuned EDCA can provide a delay less than 5 ms with the same efficiency. Thus, the area where the Smart PCA gains in efficiency for $T_{period} = 20$ ms is negligible. It happens because such traffic periods and high M values result in high intensity of traffic; thus, there is high contention between RTA STAs during the reservations.

The results for $T_{period} = 50$ ms with the same values of $\sigma = 10$ μ s and $M = 5$ are shown in Figure 6b. In this case, the size of the area where Smart PCA provides the given quantile with the best efficiency becomes significant due to lower intensity and hence less frequent contention between RTA STAs. So, if the delay requirement lies between

3.2 ms and 4.6 ms, the efficiency gain reaches 100% compared to the Tuned EDCA and 30% compared to PCA. For stricter requirements, the efficiency for PCA and Smart PCA are equal because the TXOP Limit value is set low and condition (1) is met, so the methods have the same PCA procedure.

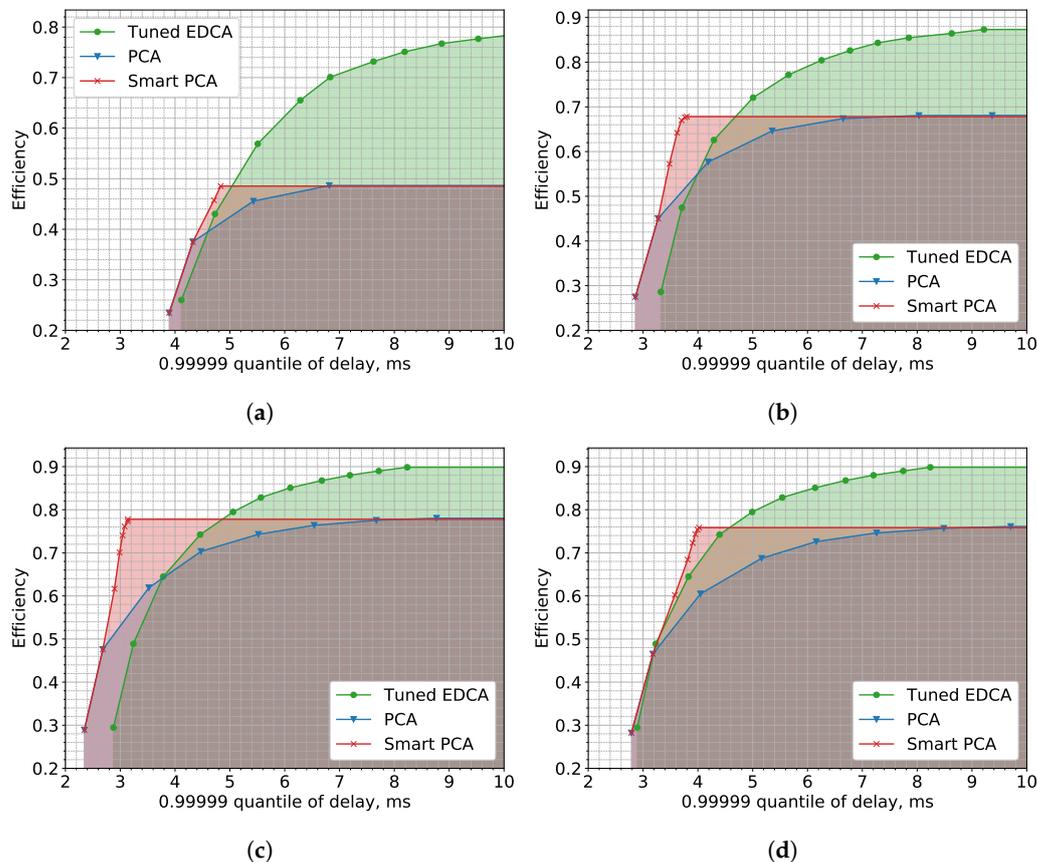


Figure 6. Dependences of the maximum efficiency on the provided delay quantile (part I). (a) $T_{period} = 20\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 5$. (b) $T_{period} = 50\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 5$. (c) $T_{period} = 100\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 5$. (d) $T_{period} = 100\text{ ms}$, $\sigma = 100\ \mu\text{s}$, $M = 5$.

Let us increase the period to $T_{period} = 100\text{ ms}$ (Figure 6c). The area of delay requirements for which the Smart PCA is most effective becomes wider for the same reason as above. Efficiency gain compared to Tuned EDCA reaches 100% ($\approx 70\%$ against $\approx 36\%$ for 3 ms delay limit) and 40% compared to PCA ($\approx 78\%$ against $\approx 56\%$ for 3.2 ms delay limit). As the period of RTA traffic increases, the efficiency advantage of the Smart PCA grows, because RTA transmissions consume channel time less often.

Figure 6d shows the results for the same period $T_{period} = 100\text{ ms}$ and the RTA STAs number $M = 5$, but for increased $\sigma = 100\ \mu\text{s}$. The Smart PCA guarantees a delay of 4 ms with maximum efficiency of 76%, while the Tuned EDCA provides the same delay with an efficiency of 68%. Thus, with a large value of σ , the efficiency gain of the Smart PCA drops from about 100% to 10%. The reason is that an increase in σ lengthens the interval T_{exp} , due to which the PCA procedure consumes more channel time.

Let us reduce the number of RTA STAs to $M = 3$ for $T_{period} = 20\text{ ms}$ and $\sigma = 10\ \mu\text{s}$ (Figure 7a). Compared to the scenario with $M = 5$, the area where the Smart PCA shows the best efficiency increases. At the same time, the gain in efficiency compared to the Tuned EDCA reaches values of more than 100%, and compared to the PCA, it reaches more than 30%. Thus, in scenarios with a small number of RTA STAs, the developed method is the most efficient under strict delay requirements for RTA traffic with a period of $T_{period} = 20\text{ ms}$. It is explained by the fact that a decrease in the number of RTA STAs

reduces the total consumption of the channel time and the delay associated with contention between RTA STAs.

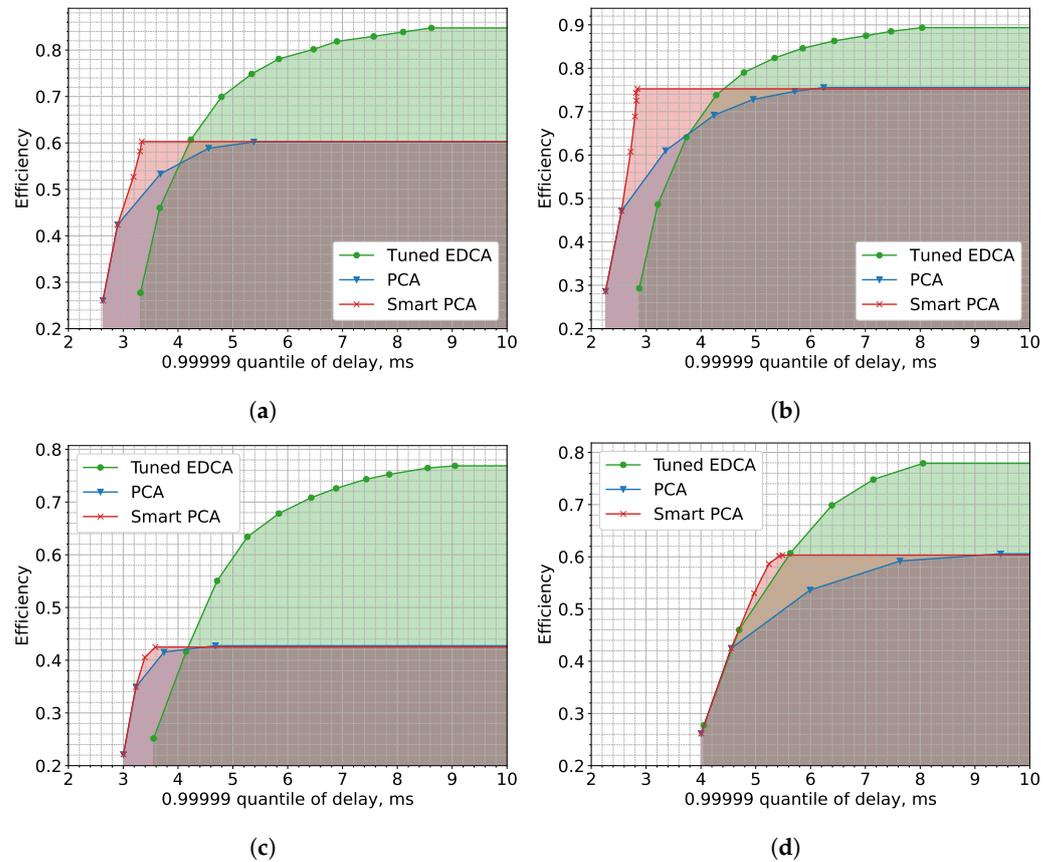


Figure 7. Dependences of the maximum efficiency on the provided delay quantile (part II). (a) $T_{period} = 20\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 3$. (b) $T_{period} = 50\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 3$. (c) $T_{period} = 10\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 3$. (d) $T_{period} = 100\text{ ms}$, $\sigma = 10\ \mu\text{s}$, $M = 15$.

The results for $T_{period} = 50\text{ ms}$ with the same values of $\sigma = 10\ \mu\text{s}$ and $M = 3$ are shown in Figure 7b. In contrast to the scenario with $M = 5$, the maximum efficiency gain of the Smart PCA increases to 150% compared with the Tuned EDCA and to 40% compared with the PCA. This result shows once again that with fewer RTA STAs, the Smart PCA method developed in the paper significantly improves channel usage efficiency.

Let us consider more intensive RTA traffic. Figure 7c shows results for $T_{period} = 10\text{ ms}$ and $M = 3$. With such intensive traffic, RTA STAs consume more channel time, which significantly reduces efficiency for PCA and Smart PCA methods because of reservation time. At the same time, a small number of RTA STAs leads to a low probability of transmission in a reservation of another STA, so the Smart PCA gain in comparison with the PCA is negligible.

In case of low intensity of RTA traffic ($T_{period} = 100\text{ ms}$) but a large number of RTA STAs ($M = 15$), we obtain an opposite result (Figure 7d). The Smart PCA has a significant gain in comparison with PCA because more RTA STAs can transmit in an alien reservation. However, high M also increases the contention between RTA STAs during channel reservations, which makes it almost unprofitable to use the Smart PCA in comparison with Tuned EDCA.

Finally, we consider non-strict requirements for delay (5–10 ms). In all the above scenarios, the Tuned EDCA is the most effective in this area of delay limits. For example, the Tuned EDCA can meet the delay limit of 10 ms with high efficiency (up to 80–90%). Thus,

the Smart PCA is more efficient than the Tuned EDCA only in case of strict requirements for the transmission delay of RTA frames.

6. Conclusions

In this paper, we propose smart preliminary channel access (Smart PCA) for serving real-time traffic in Wi-Fi networks. We compare the developed method with two previously proposed: PCA and limiting the time for which STAs can occupy the channel for non-urgent transmissions. The comparison takes into account both the QoS for RTA traffic and the channel usage efficiency for STAs with delay-insensitive traffic.

We show through simulations that the Smart PCA can provide significantly lower delay than the PCA in the presence of multiple RTA STAs in the network. The Smart PCA method proposed in this paper is the most effective in the case of strict requirements for maximum delay (<3 ms) and low intensity of RTA traffic. At the same time, an increase in the deviation of the data arrival time reduces the efficiency of the Smart PCA, so this method should be used only in the case of predictable RTA traffic.

Author Contributions: Conceptualization, D.B., E.K., and A.L.; formal analysis, D.B. and E.K.; investigation, K.C. and D.B.; methodology, E.K. and A.L.; software, K.C. and D.B.; validation, E.K. and A.L.; writing—original draft, K.C.; writing—review and editing, D.B., E.K., and A.L. All authors read and agreed to the published version of the manuscript.

Funding: The research has been carried out at IITP RAS and supported by the Russian Science Foundation (Grant No 21-19-00846, <https://rscf.ru/en/project/21-19-00846/> (accessed on 8 September 2022)).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AC	Access Category
ACK	Acknowledgement
AIFS	Arbitration Inter-Frame Space
AP	Access Point
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear-to-Send
EDCA	Enhanced Distributed Channel Access
MCS	Modulation and Coding Scheme
MLO	Multi-Link Operation
MPDU	MAC Protocol Data Unit
NAV	Network Allocation Vector
OFDMA	Orthogonal Frequency Division Multiple Access
PCA	Preliminary Channel Access
PHY	Physical layer
QoS	Quality of Service
QTP	Quiet Time Period
RTA	Real-Time Application
RTS	Request-to-Send
r-TWT	Restricted Target Wake Time
SIFS	Short Inter-Frame Space
STA	Station

References

1. Cavalcanti, D.; Bush, S.; Illouz, M.; Kronauer, G.; Regev, A.; Venkatesan, G. Wireless TSN—Definitions, Use Cases & Standards Roadmap. *Avnu Alliance* **2020**, 1–16.
2. Li, Z.; Uusitalo, M.A.; Shariatmadari, H.; Singh, B. 5G URLLC: Design Challenges and System Concepts. In Proceedings of the 2018 15th International Symposium on Wireless Communication Systems (ISWCS), Lisbon, Portugal, 28–31 August 2018; pp. 1–6. [[CrossRef](#)]
3. Rao, J.; Vrzic, S. Packet Duplication for URLLC in 5G: Architectural Enhancements and Performance Analysis. *IEEE Netw.* **2018**, *32*, 32–40. [[CrossRef](#)]
4. Karamyshev, A.; Khorov, E.; Krasilov, A.; Akyildiz, I.F. Fast and accurate analytical tools to estimate network capacity for URLLC in 5G systems. *Comput. Netw.* **2020**, *178*, 107331. [[CrossRef](#)]
5. Khorov, E.; Krasilov, A.; Selnitskiy, I.; Akyildiz, I.F. A framework to maximize the capacity of 5g systems for ultra-reliable low-latency communications. *IEEE Trans. Mob. Comput.* **2020**, *20*, 2111–2123. [[CrossRef](#)]
6. Ali, R.; Zikria, Y.B.; Bashir, A.K.; Garg, S.; Kim, H.S. URLLC for 5G and Beyond: Requirements, Enabling Incumbent Technologies and Network Intelligence. *IEEE Access* **2021**, *9*, 67064–67095. [[CrossRef](#)]
7. IEEE 802.11 Real Time Applications TIG Report. Available online: <https://mentor.ieee.org/802.11/dcn/18/11-18-2009-06-0rta-rta-report-draft.docx>, (accessed on 29 August 2022).
8. Avdotin, E.; Bankov, D.; Khorov, E.; Lyakhov, A. Resource Allocation Strategies for Real-Time Applications in Wi-Fi 7. In Proceedings of the 2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Odessa, Ukraine, 26–29 May 2020; pp. 1–6. [[CrossRef](#)]
9. Yang, M.; Li, B.; Yan, Z.; Yan, Y. AP Coordination and Full-duplex enabled Multi-band Operation for the Next Generation WLAN: IEEE 802.11be (EHT). In Proceedings of the 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), Xi'an, China, 23–25 October 2019; pp. 1–7. [[CrossRef](#)]
10. Naik, G.; Ogbe, D.; Park, J.M.J. Can Wi-Fi 7 Support Real-Time Applications? On the Impact of Multi Link Aggregation on Latency. In Proceedings of the ICC 2021—IEEE International Conference on Communications, Montreal, QC, Canada, 14–23 June 2021; pp. 1–6. [[CrossRef](#)]
11. Carrascosa, M.; Geraci, G.; Knightly, E.; Bellalta, B. An Experimental Study of Latency for IEEE 802.11be Multi-link Operation. In Proceedings of the ICC 2022—IEEE International Conference on Communications, Seoul, Korea, 16–20 May 2022; pp. 2507–2512. [[CrossRef](#)]
12. Park, H.; You, C. Latency Impact for Massive Real-Time Applications on Multi Link Operation. In Proceedings of the 2021 IEEE Region 10 Symposium (TENSYP), Jeju, Korea, 23–25 August 2021; pp. 1–5. [[CrossRef](#)]
13. Genc, E.; Del Carpio, L.F. Wi-Fi QoS Enhancements for Downlink Operations in Industrial Automation Using TSN. In Proceedings of the 2019 15th IEEE International Workshop on Factory Communication Systems (WFCS), Sundsvall, Sweden, 27–29 May 2019; pp. 1–6. [[CrossRef](#)]
14. *IEEE P802.11beTM/D2.0 Draft Standard for Information Technology — Telecommunications and Information Exchange between Systems Local and Metropolitan Area Networks — Specific Requirements*; IEEE: Piscataway, NJ, USA, 2022.
15. Bankov, D.; Chemrov, K.; Khorov, E. Tuning Channel Access to Enable Real-Time Applications in Wi-Fi 7. In Proceedings of the 2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Brno, Czech Republic, 5–7 October 2020; pp. 20–25. [[CrossRef](#)]
16. Bellalta, B. IEEE 802.11 ax: High-efficiency WLANs. *IEEE Wirel. Commun.* **2016**, *23*, 38–46. [[CrossRef](#)]
17. Deng, D.J.; Lin, Y.P.; Yang, X.; Zhu, J.; Li, Y.B.; Luo, J.; Chen, K.C. IEEE 802.11 ax: Highly efficient WLANs for intelligent information infrastructure. *IEEE Commun. Mag.* **2017**, *55*, 52–59. [[CrossRef](#)]
18. Lee, K.h. Performance analysis of the IEEE 802.11 ax MAC protocol for heterogeneous Wi-Fi networks in non-saturated conditions. *Sensors* **2019**, *19*, 1540. [[CrossRef](#)] [[PubMed](#)]
19. Naribole, S.; Lee, W.B.; Ranganath, A. Impact of MU EDCA Channel Access on IEEE 802.11ax WLANs. In Proceedings of the 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, HI, USA, 22–25 September 2019; pp. 1–5. [[CrossRef](#)]
20. Bhattarai, S.; Naik, G.; Park, J.M.J. Uplink resource allocation in IEEE 802.11 ax. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6. [[CrossRef](#)]
21. Rathor, R.G.; Joshi, R.D. Performance Analysis of IEEE802.11ax (Wi-Fi 6) Technology using Multi-user MIMO and Up-Link OFDMA for Dense Environment. In Proceedings of the 2021 IEEE 2nd International Conference on Applied Electromagnetics, Signal Processing, & Communication (AESPC), Bhubaneswar, India, 26–28 November 2021; pp. 1–7. [[CrossRef](#)]
22. Islam, G.Z.; Kashem, M.A. A Proportional Scheduling Protocol for the OFDMA-Based Future Wi-Fi Network. *J. Commun.* **2022**, *17*, 322–338. [[CrossRef](#)]
23. Uwai, T.; Miyamoto, T.; Nagao, Y.; Lanante, L.; Kurosaki, M.; Ochi, H. Adaptive backoff mechanism for OFDMA random access with finite service period in IEEE802. 11ax. In Proceedings of the 2016 IEEE Conference on Standards for Communications and Networking (CSCN), Berlin, Germany, 31 October–2 November 2016; pp. 1–6. [[CrossRef](#)]
24. Bai, J.; Fang, H.; Suh, J.; Aboul-Magd, O.; Au, E.; Wang, X. Adaptive uplink OFDMA random access grouping scheme for ultra-dense networks in IEEE 802.11 ax. In Proceedings of the 2018 IEEE/CIC International Conference on Communications in China (ICCC), Beijing, China, 16–18 August 2018; pp. 34–39. [[CrossRef](#)]

25. Wang, J.; Wu, M.; Chen, Q.; Zheng, Y.; Zhu, Y.h. Probability Complementary Transmission Scheme for Uplink OFDMA-based Random Access in 802.11ax WLAN. In Proceedings of the 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 15–18 April 2019; pp. 1–7. [CrossRef]
26. Kim, Y.; Kwon, L.; Park, E.C. OFDMA backoff control scheme for improving channel efficiency in the dynamic network environment of IEEE 802.11 ax WLANs. *Sensors* **2021**, *21*, 5111. [CrossRef]
27. Cheng, R.G.; Yang, C.M.; Firmansyah, B.S.; Harwahyu, R. Uplink OFDMA-Based Random Access Mechanism With Bursty Arrivals for IEEE 802.11 ax Systems. *IEEE Netw. Lett.* **2021**, *4*, 34–38. [CrossRef]
28. Li, J.; Bose, S.K.; Shen, G. Cooperative Resource Scheduling for Time-sensitive Services in an Integrated XGS-PON and Wi-Fi 6 Network. *IEEE Commun. Lett.* **2022**, *26*, 1338–1342. [CrossRef]
29. Kosek-Szott, K.; Domino, K. An Efficient Backoff Procedure for IEEE 802.11ax Uplink OFDMA-Based Random Access. *IEEE Access* **2022**, *10*, 8855–8863. [CrossRef]
30. Abramson, N. The ALOHA system: Another alternative for computer communications. In Proceedings of the Fall Joint Computer Conference, Houston TX, USA, 17–19 November 1970; pp. 281–285.
31. Laya, A.; Kalalas, C.; Vazquez-Gallego, F.; Alonso, L.; Alonso-Zarate, J. Goodbye, ALOHA! *IEEE Access* **2016**, *4*, 2029–2044. [CrossRef]
32. Kim, Y.; Oh, S.; Kim, G.; Jeong, J. Performance Analysis of QTP-based S2S Transmission in IEEE 802.11ax WLANs. In Proceedings of the 2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), Jeju Island, Korea, 17–20 August 2021; pp. 280–282. [CrossRef]
33. Khorov, E.; Lyakhov, A.; Ivanov, A.; Akyildiz, I.F. Modeling of Real-Time Multimedia Streaming in Wi-Fi Networks With Periodic Reservations. *IEEE Access* **2020**, *8*, 55633–55653. [CrossRef]
34. López-Pérez, D.; Garcia-Rodriguez, A.; Galati-Giordano, L.; Kasslin, M.; Doppler, K. IEEE 802.11 be extremely high throughput: The next generation of Wi-Fi technology beyond 802.11 ax. *IEEE Commun. Mag.* **2019**, *57*, 113–119. [CrossRef]
35. Khorov, E.; Levitsky, I.; Akyildiz, I.F. Current Status and Directions of IEEE 802.11be, the Future Wi-Fi 7. *IEEE Access* **2020**, *8*, 88664–88688. [CrossRef]
36. Deng, C.; Fang, X.; Han, X.; Wang, X.; Yan, L.; He, R.; Long, Y.; Guo, Y. IEEE 802.11 be Wi-Fi 7: New challenges and opportunities. *IEEE Commun. Surv. Tutorials* **2020**, *22*, 2136–2166. [CrossRef]
37. Yang, M.; Li, B. Survey and perspective on extremely high throughput (EHT) WLAN—IEEE 802.11 be. *Mob. Netw. Appl.* **2020**, *25*, 1765–1780. [CrossRef]
38. Chen, C.; Chen, X.; Das, D.; Akhmetov, D.; Cordeiro, C. Overview and Performance Evaluation of Wi-Fi 7. *IEEE Commun. Stand. Mag.* **2022**, *6*, 12–18. [CrossRef]
39. Adame, T.; Carrascosa-Zamacois, M.; Bellalta, B. Time-sensitive networking in IEEE 802.11 be: On the way to low-latency WiFi 7. *Sensors* **2021**, *21*, 4954. [CrossRef] [PubMed]
40. Bankov, D.; Khorov, E.; Lyakhov, A.; Sandal, M. Enabling real-time applications in Wi-Fi networks. *Int. J. Distrib. Sens. Netw.* **2019**, *15*, 1–11. [CrossRef]
41. Rentschler, M.; Laukemann, P. Performance analysis of parallel redundant WLAN. In Proceedings of the 2012 IEEE 17th International Conference on Emerging Technologies Factory Automation (ETFA 2012), Krakow, Poland, 17–21 September 2012; pp. 1–8. [CrossRef]
42. Cena, G.; Scanzio, S.; Valenzano, A.; Zunino, C. An enhanced MAC to increase reliability in redundant Wi-Fi networks. In Proceedings of the 2014 10th IEEE Workshop on Factory Communication Systems (WFCS 2014), Toulouse, France, 5–7 May 2014; pp. 1–10. [CrossRef]
43. Cena, G.; Scanzio, S.; Valenzano, A. Experimental evaluation of seamless redundancy applied to industrial Wi-Fi networks. *IEEE Trans. Ind. Inform.* **2017**, *13*, 856–865. [CrossRef]
44. Multi-Link for Low Latency. Available online: <https://mentor.ieee.org/802.11/dcn/20/11-20-0093-02-00be-multi-link-for-low-latency.pptx>, (accessed on 29 August 2022).
45. Lacalle, G.; Val, I.; Seijo, O.; Mendicute, M.; Cavalcanti, D.; Perez-Ramirez, J. Analysis of Latency and Reliability Improvement with Multi-Link Operation over 802.11. In Proceedings of the 2021 IEEE 19th International Conference on Industrial Informatics (INDIN), Palma de Mallorca, Spain, 21–23 July 2021; pp. 1–7. [CrossRef]
46. Carrascosa-Zamacois, M.; Galati-Giordano, L.; Jonsson, A.; Geraci, G.; Bellalta, B. Performance and Coexistence Evaluation of IEEE 802.11 be Multi-link Operation. *arXiv* **2022**, arXiv:2205.15065.
47. Reducing Channel Access Delay for RTA Traffic. Available online: <https://mentor.ieee.org/802.11/dcn/19/11-19-1960-01-00be-reducing-channel-access-delay-for-rta-traffic.pptx> accessed on (accessed on 29 August 2022).
48. ns-3 l a Discrete-Event Networks Simulator for Internet Systems. Available online: <https://www.nsnam.org/> (accessed on 27 September 2022).