# IEEE 802.11ax Uplink Scheduler to Minimize Delay: a Classic Problem with New Constraints

Dmitry Bankov, Andre Didenko, Evgeny Khorov, Vyacheslav Loginov, Andrey Lyakhov
Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia
Email: {bankov, dida, khorov, loginov, lyakhov}@iitp.ru

***Abstract*—In order to meet the continuously increasing demands for high throughput in wireless networks, IEEE 802 LAN/MAN Standard Committee is developing IEEE 802.11ax: a new amendment for the Wi-Fi standard. This amendment provides various ways to improve the efficiency of Wi-Fi. The most revolutionary one is OFDMA. Apart from obvious advantages, such as decreasing overhead for short packet transmission at high rates and improving robustness to frequency selective interference, being used for uplink transmission, OFDMA can increase power spectral density and, consequently, user data rates. However, the gain of OFDMA mainly depends on the resource scheduling between users. The peculiarities of OFDMA implementation in Wi-Fi completely change properties of classic schedulers used in other OFDMA systems, e.g. LTE. In the paper, we consider the usage of OFDMA in Wi-Fi for uplink transmission. We study peculiarities of OFDMA in Wi-Fi, adapt classic schedulers to Wi-Fi, explaining why they do not perform well. Finally we develop a novel scheduler, MUTAx, and evaluate its performance with simulation.**

***Index Terms*—Wi-Fi, IEEE 802.11ax, High Effiency WLAN, OFDMA, Scheduling**

## I. Introduction

Nowadays, Wi-Fi has become the main technology for wireless local area networks. High number of Wi-Fi devices, as well as the number of deployed networks, leads to huge interference. To improve efficiency of Wi-Fi networks in existing and emerging indoor and outdoor scenarios, Wi-Fi community is currently developing a new standard, namely IEEE 802.11ax. In contrast to 11ac, PHY layer data rate of which tenfold excels the one of the preceding technology (namely 11n), the expected quadruple growth of user throughput in 11ax networks will be caused mostly by advanced channel access techniques, rather than by PHY data rates increased by just 37%.

The main feature of 11ax is Orthogonal Frequency Division Multiple Access (OFDMA), which extends the legacy Wi-Fi Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) by introducing a possibility to divide channel resources in frequency domain. Since 11a, Wi-Fi has been using Orthogonal Frequency Division Multiplexing (OFDM). However, with OFDM, at any time instant all tones (also referred to as subcarriers) are used to transmit data for one user, while OFDMA allows assigning various tones to different users. The efficiency of OFDMA significantly depends on how the tones are scheduled between users. However the 11ax standard will provide only a flexible framework, without any predefined scheduling algorithms.

The research was supported by the Russian Science Foundation (agreement No 16-19-10687).

Fortunately, scheduling problem has been carefully investigated in cellular networks, like LTE, where OFDMA has appeared much earlier. So at first sight, it is worth to use one of the existing cellular schedulers and adapt it to Wi-Fi peculiarities. This work itself is not easy, since OFDMA fundamentals in Wi-Fi differ from that in LTE. Moreover, the features of 11ax break assumptions used to derive schedulers for LTE. Thus, nobody can guarantee that being applied to 11ax networks the LTE scheduler remain the best one.

In this paper, we compare OFDMA schemes in IEEE 802.11ax and LTE, and analyze problems that arise while developing the schedulers for 11ax networks. We also make the first step in this direction and consider a problem of minimizing the average delivery time for uplink transmission. Then we show why well-known schedulers designed for a similar problem are not optimal for 11ax networks and develop a new scheduler that outperforms them.

The rest of the paper is organized as follows. Section II briefly describes the main features of OFDMA in 802.11ax, reviews literature and states the problem. In Section III we design a brand new scheduler for 11ax networks. We prove its high efficiency in Section IV. Section V concludes the paper.

## II. OFDMA in IEEE 802.11ax Networks

### A. Main Features

In contrast to LTE with a rigid traffic-independent time-frequency numerology, in 11ax networks, OFDMA works at the per-frame bases upon native to Wi-Fi Enhanced Distributed Channel Access (EDCA), a sort of decentralized CSMA/CA joint with automatic repeat request. OFDMA frames start with a common preamble, which can be decoded by legacy devices. Having received the preamble, a STA learns the duration of the frame and considers the medium as busy till the end of the frame. The rest of the frame — understandable only by 11ax STAs — can be formed according to OFDMA concept, i.e. various tones of the frame can be assigned to different STAs.

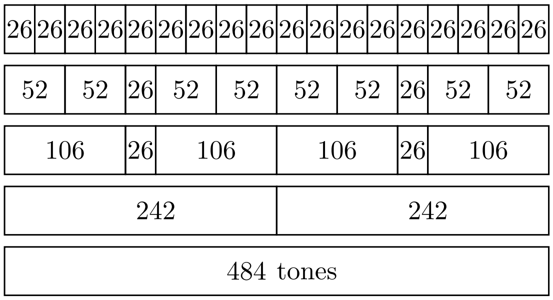


Figure 1: RU locations in 40 MHz channel

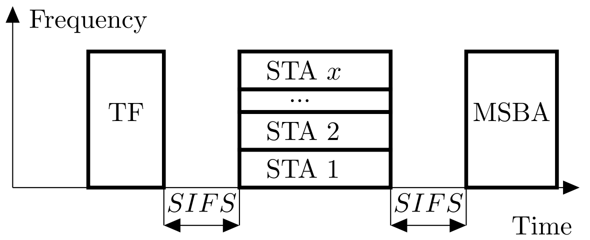


Figure 2: Frame handshake for UL OFDMA transmission.

In 11ax networks, a group of tones assigned to a STA is called Resource Unit (RU). The IEEE 802.11ax defines RUs that consist of 26, 52, 106, 242, 484, 996 and $2 \times 996$ OFDM tones. The set of available RUs depends on the channel width, e.g., in a 40 MHz channel, the STAs can use RUs of up to 484 tones. A wide RUs can be split into approximately twice narrower RUs *independently from others*. The two exceptions are as follows. 1) A 996-tone RU can be split into two 484-tone RU and one 26-tone RU. 2) A 242-tone RU can be split into two 106-tone RU and one 26 tone RU, see Fig. 1.

In LTE networks, a scheduler can allocate an arbitrary subset of downlink (DL) RUs for a user or an arbitrary *interval* of UL RUs for a user. 11ax imposes much stricter constrains on RU selection, since both in DL and UL, a STA cannot be assigned to more than one RU.

Another set of constraints limits the maximal MCS that can be used in narrow RUs and the usage of MU MIMO. Specifically, MU MIMO is only available in 106-tone and greater RUs, while the novel 1024QAM can be used only in 242-tone and greater RUs. Such constraints complicate scheduling problem in 11ax networks.

An OFDMA frame may contain RUs with different number of tones, however all RUs inside the frame must have the same duration. For that, the STAs may use novel flexible fragmentation or padding.

All transmissions inside one OFDMA frame must be synchronized, i.e. start and finish at the same time instants. This can be easily done in DL, where an OFDMA frame is generated by the Access Point (AP). However, to provide synchronization for uplink, the AP can use novel Trigger Frame (TF). SIFS after TF reception, the STAs transmit their parts of UL OFDMA frame. If needed, the AP acknowledges reception of each part by sending a set of ACKs inside an OFDMA frame or by sending Multi-STA Block Acknowledgment (MSBA).

It is the AP which, for both DL and UL transmissions, determines modulation and coding scheme (MCS), duration, RU assignment and other OFDMA parameters. Such information can be transmitted in frame headers (for DL) and in the Trigger frame (for UL). An example of UL OFDMA transmission is shown in Fig. 2.

As mentioned above, 11ax moves decision making logic from the STAs to the AP that determines which STA transmits, when, in which RU, how much data it transmits, etc. To make a correct decision, the AP needs to be aware of STAs' buffered traffic and channel conditions. This information can be periodically requested by the AP. Moreover, to notify the AP about arrived packets, a STA can proactively send so-called Buffer Status Report (BSR). For that, it can aggregate BSR with a data frame sent in the dedicated RU. Another option is legacy EDCA. Thanks to the recent change in the 11ax draft standard, the AP can separately tune EDCA parameters of scheduled STAs in such a way, that they do not contend for the channel with unscheduled STAs. It means that if a STA does not obtain UL RUs, it likely delivers a BSR at the first transmission attempt because of the extremely low contention. The third approach to send a BSR is using OFDMA Random Access. For that, the AP can allocate one or several RUs for random access and the contending STAs will randomly choose one of such RUs. Anyway, thanks to these methods, the AP can quickly obtain information about new frames buffered at the associated STAs and waiting for UL RUs.

OFDMA brings many benefits to 11ax networks. First it makes transmission more reliable to frequency selective interference and fading. This is especially important for wide 160 MHz channels introduced in 11ac. Second, with OFDMA we can glue short packets destined for or originated from various STAs, which significantly reduces overhead caused by PHY headers and channel access time and interframe spaces. Third, for UL transmissions by edge STAs it makes sense to use narrow channels instead of wide ones. Indeed, since the STAs are spatially separated they can simultaneously transmit at the maximal allowed power without breaking legal limitations on the emitted energy. Thus, by reducing RU width, we increase power spectral density received from the STAs and can use higher MCS. In other words, by allowing numerous spatially separated STAs to transmit in parallel, we increase the cumulative received power in comparison to the case when only one STA transmits. For edge STAs, increasing power spectral density leads to a higher MCS, so the average amount of data received from an OFDMA frame is higher than that from a legacy. Thus in contrast to LTE, in 802.11ax networks the rates in UL RUs are non-additive, i.e. if a STA transmits in twice wider RU, it is not guaranteed that it transmits twice more data. This effect will be carefully investigated in the paper.

### B. Related Works

Despite the fact that 802.11ax amendment is expected to be finished by 2019, it has been already studied in the literature [1]–[5].

In [2], authors study the performance of the network consisting of the legacy and 802.11ax STAs. Authors propose an approach for optimal channel access parameter values selection, which guarantees fairness between legacy and 802.11ax and significantly increases the number of the OFDMA transmissions for 802.11ax STAs. However, the model developed in [2] allows only estimating the number of the OFDMA transmissions in the network, but not the achievable data rate.

Many studies on 802.11ax performance were presented at the meetings of the 802.11ax Task Group (TGax), e.g. [4], [5]. However, although many network topologies have been already studied, the only considered scheduler in TGax has been the random scheduler with a static RU configuration.

The problem of scheduling in 11ax uplink has been considered in [3], which focuses on selection of the RU duration. The authors initially assume that the AP does not know the amount of data the STAs have and thus does not know the correct duration of RUs that should be allocated. They propose a scheme that can be used by the AP to obtain this information and derive the best way to select the duration of STAs' transmission in terms of throughput and energy consumption. However, they do not consider any specific way to divide the channel into RUs and to assign the RUs to the STAs. Neither do they consider random access and the possibility of STAs to deliver RSRs by aggregating them to the transmitted data.

Random access for UL OFDMA in 802.11ax has been studied in [6]. The authors consider a scenario, when STAs transmit saturated data flows and use only random access for transmission. They describe a mathematical model of transmission with OFDMA random access and use it for analysis and optimization of the network performance in terms of throughput and percentage of successful RUs. The methodology introduced in the paper might be used to tune the random access for OFDMA UL, however, usage of deterministic access is more feasible in terms of channel efficiency for the studied scenario.

The problem of resource allocation in cellular networks is widely studied in literature. In many papers, it is formulated as the following optimization problem. Consider a network with a base station (BS) and $N$ associated users. Each time unit, the BS runs scheduler which allocates $M$ RUs to some STAs in order to maximize some network-wide utility function. The most widely used utility functions and corresponding schedulers are listed below.

To maximize cumulative throughput $S = \sum\limits_{i=1}^{N} S_i(t)$ at time instant $t$, the BS uses the Max Rate (MR) scheduler, which considers RUs one by one and assigns each RU to a user with the highest nominal data rate $r_i^j$ in this RU. However, as shown in many papers (e.g. [7]), under high load the MR scheduler blocks users with low instant rates.

To avoid this problem, the Proportional Fair (PF) scheduler has been proposed. It maximizes $\sum_{i=1}^{N} \log S_i(t)$. For that, RU $j$ shall be allocated to user $\hat{i}_{PF} = \arg\max_i \frac{r_i^{(j)}(t)}{S_i(t-1)}$. Authors of [8] show that in long term the PF scheduler gives the same amount of channel time for all users, therefore resulting in fair resource allocation.

Both MR and PF have been designed for saturated (infinite) flows. From the practical point of view, it is worth to consider unsaturated scenario with finite flows of given length. Such flows can represent HTTP requests or files transmitted in a wireless network. For such traffic, it is worth to consider upload time which affects user perceived quality of experience (QoE). If the channel properties do not change with time, the rate in different RUs is the same and additive, the Shortest Remaining Time First (SRTF) scheduler is proven to provide minimal average upload time. Simplistically, SRTF allocates all RUs to a user for which $\frac{D_i(t)}{r_i}$ is the minimal one, where $D_i(t)$ is the remaining amount of data of user $i$ and $r_i$ is its rate in the whole channel. As explained in Section II-A, the assumptions made to derive such a scheduler are not met for UL OFDMA transmission in 11ax networks, which brings us to the following problem statement.

### C. Problem Statement

Consider a scenario with an AP and a group of STAs associated to it. From time to time, STAs generate finite data flows to be delivered to the AP. To notify the AP about new flow, the STAs use EDCA to transmit BSR. Thus the whole channel can be allocated to those STAs, which are known by the AP as having traffic.

In the paper, we consider the following problem: *to design a scheduler for UL OFDMA data transmission in 11ax networks that minimizes the average upload time.*

## III. Design of the Scheduler

In this section, we design a novel scheduler called MUTAX, Minimizing Upload Time in 11AX networks. While deriving formulae we neglect the effects related to packetization (including aggregation and fragmentation) overhead. Apart from that, for shortness we consider only $n$ STAs with flows and assume that each STA has only one flow. Both STA and flow are denoted as $i, i = 1, ..., n$.

Let slot be a time interval between two consequent TFs. It should be noted that slots may have different duration. The maximal one is related to the standard limit of 5484 µs for the physical protocol data unit (PPDU) duration. A slot can be shorter, if at least one STA which transmits in this slot has no more data.

The MUTAX algorithm has two steps. At the first step, we order existing flows and calculate the sum upload time of the flows, as if we used exhaustive service. At the second stage, we try to improve the sum upload time by serving some flows in parallel. Let us consider the steps in detail.

At the first step, we serve each flows exhaustively. Except for waiting, the time needed to upload flow $i$ equals $t_i = \frac{D_i}{r_i}$. The first STA finishes delivering its flow by time $t_1$. The second STA starts right after the first one and delivers its flow by time $t_2 + t_1$, etc. As the result, the total upload time for existing flows is

$$T_{step1} = \sum_{i=1}^{n} (n-i)\, t_i.$$

Obviously, to minimize the sum upload time we have to sort the STAs in the ascending order by $t_i$.

At the second step, we divide the channel into several RUs. Let $m$ be the number of RUs and $j, 1 \leq j \leq m$ be the index of RU in the considered RU configuration. $x_i^j$ is an indicator which equals 1 if STA $i$ is assigned to RU $j$, and 0, otherwise. For shortness, $X$ is the two dimensional matrix of $\left\{x_i^j\right\}$ representing RU assignment to the STAs.

With defined notation, the total upload time $T\left(X\right)$ of existing flows differs from $T_{step1}$ in the following way. First, the upload time of each of $n$ flows increases by the current slot duration $\tau$. Second, if $x_i^j = 1$, the remaining amount of data of flow $i$ decreases by the amount of data the STA transmits in RU $j$ of the current slot: $\Delta D_i^j = \min\left\{D_i, \tau \times r_i^j\right\}$. Thus

$$T\left(X\right) = n\tau + \sum_{i=1}^{n} (n-i)\, \frac{D_i - \sum_{k=1}^{m} x_i^j \Delta D_i^j}{r_i}$$

Since both $\tau$ and $\Delta D_i^j$ depend on the resource allocation $X$, minimization of $T(X)$ requires exhaustive search by possible ways to allocate the RUs to the STAs, and to simplify the task we propose an heuristic approach based on two assumptions. Firstly, we neglect the change of $n\tau_t$ for different allocations, as a slot cannot be too long due to standard limitations. Secondly, we sort the STAs in the ascending order by $t_n$ only once, before considering different ways to assign RUs to the STAs. Under these assumptions, to minimize $T(X)$, we have to maximize the following expression $\sum_{k=1}^{m} \sum_{i=1}^{n} (n-i)\, \frac{x_i^j \Delta D_i^j}{r_i}$.

Let us denote $\lambda_i^j = (n-i)\, \frac{\Delta D_i^j}{r_i}$ and define the following optimization problem:

$$\begin{aligned} \max & \sum_i \sum_j x_i^j \lambda_i^j \\ \text{subject to} & \sum_i x_i^j \leq 1, \quad \forall j \\ & \sum_j x_i^j \leq 1, \quad \forall i \\ & \sum_i \sum_j x_i^j \leq m \end{aligned}$$

This problem[1] is known as the assignment problem which can be solved in polynomial time using the Kuhn-Munkres algorithm [9]. Its solution is the assignment $\hat{X}$ that yields minimal $T(X)$.

Note that assignment $\hat{X}$ is found for a specific configuration of RUs. To minimize upload time, we consider different ways to split the channel into RUs. For each configuration of RUs we solve the optimization problem and thus find the best assignment of the channel resources. The examination of RU configuration can be hastened by excluding configurations that are obviously worse than the known alternatives, e.g., if we have two STAs and a 20 MHz channel, we do consider configuration with one 106-tone RU, one 52-tone RU and three 26-tone RUs, but exclude configuration with one 106-tone RU and five 26-tone RUs.

## IV. Numerical Results

To evaluate the developed scheduler, we use the well-known NS-3 simulator [10]. We have implemented the MUTAX scheduler, as well as 11ax adaptations of PF, MR and SRTF. Since the performance of PF significantly depends on how the channel is divided into RUs, we find the best RU configuration with exhaustive search.

We run simulation in a scenario described in Section II-C. The network operates in a 40 MHz channel at 5 GHz. The flow sizes are drawn from truncated lognormal distribution with minimal, average and maximal values of 1 KB, 500 KB, and 5 MB, respectively. When a flow is delivered, the next flow is generated after a random delay drawn from truncated exponential distribution with minimal, average and maximal values of 0.1 s, 0.3 s and 0.6 s, respectively.

In first set of experiments, the STAs are located uniformly within a small circle of radius $R = 5$ m around the AP. In such a case, the channel quality is so high that the STAs use the maximal MCS to transmit in RUs of any width, except for 26-, 52- and 106-tone RUs, for which the usage of two highest MCSs is prohibited. Obviously, in this case OFDMA cannot bring any profit against SRTF, as the division of the channel without changing the MCS cannot increase the data rate, provided that the MCS can cope with the noise. This result is supported by simulation, see Fig 3, which shows that if the channel for all STAs is perfect, MUTAX yields the same upload time, as SRTF, and they both outperform the PF and MR schedulers by up to 30%. As we consider traffic model of a closed loop system, along with the reduction of upload time MUTAX provides the increase of goodput compared to the other schedulers.

The second set of experiments corresponds to the case when STAs are located in a larger circle of radius $R =$

[1]If we define $\lambda_i^j = \frac{r_i^j}{S_i}$, where $S_i$ is the amount of data transmitted by STA $i$, we obtain the optimization problem for the adaptation of the PF scheduler to 11ax.

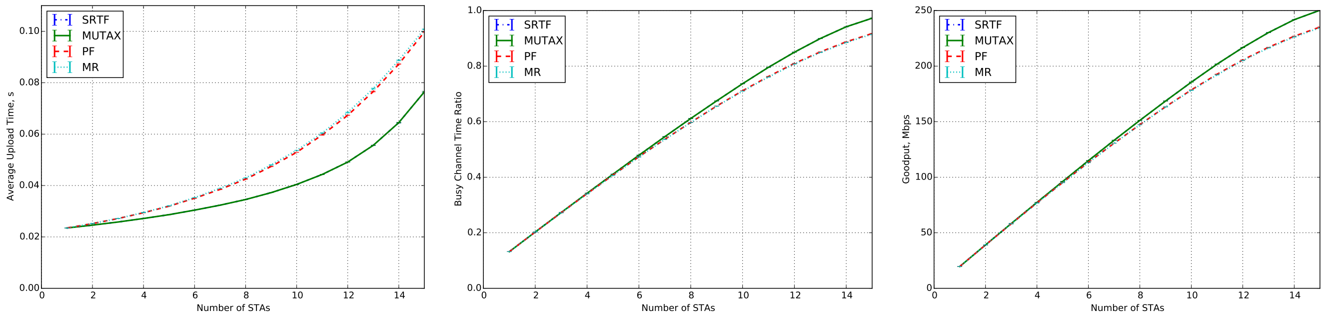


Figure 3: Upload time, busy channel time ratio and goodput vs the number of STAs in the small circle.

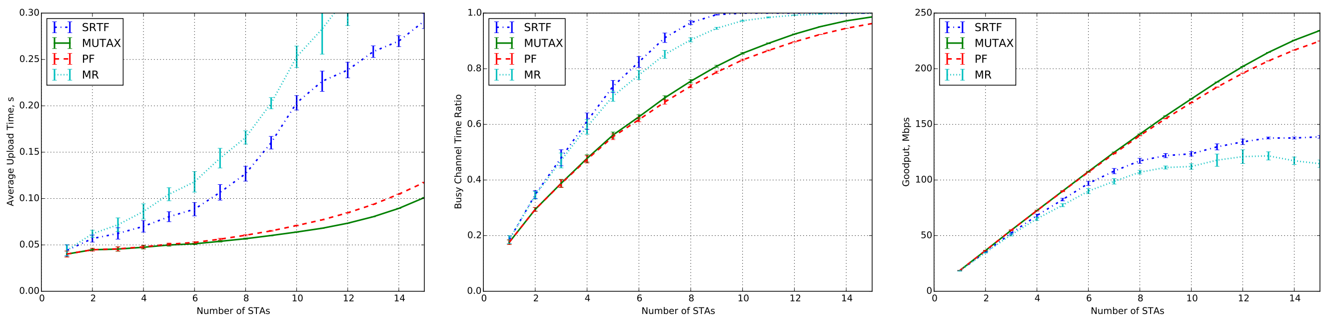


Figure 4: Upload time, busy channel time ratio and goodput vs the number of STAs in the large circle.

20 m. Such conditions provide variety of MCSs among STAs and among RU widths, so it becomes feasible to split the channel between different users. According to simulation results, see Fig. 4, in this case designed to minimize upload time the classic SRTF work much worse than even 11ax adaptation of PF. At the same time, MUTAX shows 20% lower upload time than PF. The results show that in this case the schedulers based on the exhaustive service (MR and SRTF) are much less efficient than channel-splitting schedulers (MUTAX and PF) and the gap between them increases with the number of client STAs. In a large circle scenario the gain in goodput of MUTAX against SRTF and MR is almost 100%.

## V. Conclusion

In the paper, we have studied scheduling problem in IEEE 802.11ax networks, the standard of which is currently under development. We show that because of 11ax OFDMA peculiarities the existing schedulers cannot be directly applied to the new technology. Specifically, in the paper we have considered the problem of uplink scheduling which aims to minimize average upload time in a scenario with high number of active users. We showed that depending on the scenario sometimes it is worth to use classic SRTF scheduler, while in other cases the channel should be split between several STAs in order to minimize upload time. We develop a novel scheduler, called MUTAX, which adaptively chooses the best strategy and significantly outperforms existing popular solutions.

## References

[1] B. Bellalta, "IEEE 802.11 ax: high-efficiency WLANs," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 38–46, 2016.

[2] E. Khorov, V. Loginov, and A. Lyakhov, "Several edca parameter sets for improving channel access in ieee 802.11 ax networks," in *Wireless Communication Systems (ISWCS), 2016 International Symposium on.* IEEE, 2016, pp. 419–423.

[3] M. Karaca, S. Bastani, B. E. Priyanto, M. Safavi, and B. Landfeldt, "Resource management for OFDMA based next generation 802.11 WLANs," in *Wireless and Mobile Networking Conference (WMNC), 2016 9th IFIP.* IEEE, 2016, pp. 57–64.

[4] S. Kim, H. Park, J. Kim, K. Ryu, and H. Cho, *11ax PAR Verification through OFDMA*, January 2017. [Online]. Available: https://mentor.ieee.org/802.11/dcn/16/11-16-1363-01-00ax-11ax-par-verification-through-ofdma.ppt

[5] H. Su, J. Pang, J. Zhu, J. Zhang, and C. Ma, *Preliminary 11ax PAR Verification*, September 2016. [Online]. Available: https://mentor.ieee.org/802.11/dcn/16/11-16-1363-01-00ax-11ax-par-verification-through-ofdma.ppt

[6] L. Lanante, H. Ochi, T. Uwai, Y. Nagao, M. Kurosaki, and C. Ghosh, "Performance Analysis of the 802.11ax UL OFDMA Random Access Protocol in Dense Networks," in *Communications, 2017. ICC'17. IEEE International Conference on.* IEEE, 2017.

[7] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: Key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[8] R. Kwan, C. Leung, and J. Zhang, "Proportional fair multiuser scheduling in lte," *IEEE Signal Processing Letters*, vol. 16, no. 6, pp. 461–464, 2009.

[9] F. Bourgeois and J.-C. Lassalle, "An extension of the munkres algorithm for the assignment problem to rectangular matrices," *Communications of the ACM*, vol. 14, no. 12, pp. 802–804, 1971.

[10] The ns-3 Network Simulator. [Online]. Available: http://www.nsnam.org/