ORIGINAL PAPER

WiMCA: multi-indicator client association in software-defined Wi-Fi networks

Blas Gómez[1] · Estefanía Coronado[2] · José M. Villalón[1] · Roberto Riggio[3] · Antonio Garrido[1]

Abstract

In a world with increasing traffic demands, wireless technologies aim to meet them by means of new Radio Access Technologies that provide faster connectivity. Such is the case of 4G and 5G. However, in indoor scenarios, where the capabilities of these technologies are significantly affected by the distance to the base station and the materials used in the construction of buildings, Wi-Fi is still the technology of reference thanks to its low cost and easy deployment. In this context, it is usual to find multi-AP Wi-Fi networks whose deployment has been carefully planned. However, the user-AP association decision procedure is not defined by the IEEE 802.11 standard. As a result, vendors choose selfish approaches based on signal strength. This leads to uneven user distributions and nonoptimal resource utilization. To deal with this, densification has been used over the years, but this is expensive as it needs more infrastructure. Moreover, this results in more APs in the same collision domain. To avoid the need for densification, in this paper we introduce WiMCA, a joint SDN-based user association and channel assignment solution for Wi-Fi networks that considers signal strength, channel occupancy and AP load to make better association decisions. Experimental results have demonstrated that, in terms of aggregated goodput, WiMCA outperforms approaches based on signal strength by 55%, providing better user level fairness and accommodating more users and traffic before reaching the point at which densification is needed.

Keywords Wi-Fi · SDN · IEEE 802.11 · Mobility management · Load balancing · User association

✉ Blas Gómez
Blas.Gomez@uclm.es

Estefanía Coronado
estefania.coronado@i2cat.net

José M. Villalón
JoseMiguel.Villalon@uclm.es

Roberto Riggio
roberto.riggio@ri.se

Antonio Garrido
Antonio.Garrido@uclm.es

[1] High-Performance Networks and Architectures (RAAP), University of Castilla-La Mancha, Albacete, Spain

[2] i2CAT Foundation, Barcelona, Spain

[3] RISE Research Institutes of Sweden AB, Stockholm, Sweden

1 Introduction

Our society is moving fast towards a wireless world. More and more portable devices incorporate at least one Radio Access Technology (RAT). 4G and 5G RATs provide bigger coverage areas and faster connectivity to cope with the demands of mobile traffic. However, the poor indoor signal quality caused by the distance from a base station and the materials used in the construction of buildings has led to Wi-Fi being the chosen technology for indoor scenarios. This choice is also supported by its low cost and easy deployment. Nevertheless, its contention-based scheme and its unplanned nature lead to sub-optimal performance. In addition, Wi-Fi operates on unlicensed frequency bands. This contributes to its low cost and ease of deployment, but also makes it more vulnerable to interference from co-located deployments due to the small number of available channels (especially in the 2.4 GHz band). As a result, avoiding collision domains is a difficult task. When multiple APs are in the same collision domain,

throughput degradation is to be expected, especially as the number of APs per unit area increases. Moreover, in deployments with a high number of users where a single AP cannot serve the needs of all of them, the most popular solution is densification, i.e., deploying more APs and decreasing transmission power to offload the overloaded APs, as shown in Fig. 1. Nevertheless, densification aggravates the aforementioned problems and increases infrastructure costs, thus calling for more sophisticated approaches to address this issue.

Regardless of the degree of density, user and resource management policies at the air interface may often be inefficient due to the lack of standard user-AP association methods. The IEEE 802.11 standard does not impose any specific procedure and leaves the decision to the vendor. For simplicity, the majority of vendors follow a client-driven (CD) approach, also refered to as an RSSI-based scheme in the bibliography. In this approach, stations select the AP with the highest Received Signal Strength Indicator (RSSI). However, the problem with just taking signal strength into account is that valuable information about noise and interference in the wireless medium, together with the current load of the network, is disregarded. This may lead to an uneven load distribution over the infrastructure, which contributes to suboptimal network resource allocation. This problem becomes even worse if we consider high user concentrations caused by flash events such as a boarding call to one of the gates of an airport, or an autograph session in a mall, where all clients try to connect to a small set of APs, leaving the rest idle. Studies such as [1, 2] have demonstrated that the user distribution is uneven in many scenarios such as airports or office buildings. As a result, some APs might be overloaded whereas others might be idle. As mentioned above, traditionally the solution for an overloaded AP has been densification. However, since the load distribution in the network is uneven it is clear that there are free resources that could be used to avoid the need for installing more

infrastructure. A better solution is to efficiently reshape the user distribution map, taking away a portion of the traffic from the saturated APs to enable the use of already available network resources whenever this is possible. This approach would avoid resorting to costly and difficult to manage densification solutions, relegating them to a scenario of a big growth in the number of users all over the network.

In recent years, a new paradigm has emerged as a way to redesign network functions: Software-Defined Networking (SDN) [3]. This decouples the data plane from the control plane by providing high-level programming abstractions, thus making it possible to concentrate information about the whole network in the hands of a centralized controller. This change of perspective gives the controller a privileged position that enables more advanced and precise management operations with a two-way flow of information: on the one hand, the controller gathers information in real-time from the network devices, which becomes valuable in order to proactively react in a wide range of required actions, e.g., mobility management, fault detection and remedy, etc; on the other hand, the controller can specifically define how the network should behave at a given moment in time on the basis of well-informed decisions from the aforementioned data. SDN is a consolidated paradigm and several solutions have been proposed for both wired [4] and wireless [5, 6] networks. OpenFlow [4] works as a de facto standard for SDN in wired networks. However, there is not such a clear dominance in wireless networks. Two possible frameworks for wireless networks are 5G-EmPOWER [5] or Odin [6].

In traditional IEEE 802.11 WLANs, there is no coordination between the APs that form the network. As a consequence, the information about an AP's load and available resources is not available to the others. A centralized association scheme with a general view of the network such as the one provided by an SDN-controller is able to take more efficient association decisions. Moreover,

Fig. 1 An example of densification

Overloaded AP

Densification is used

The APs with a better RSSI from the new AP connect to it
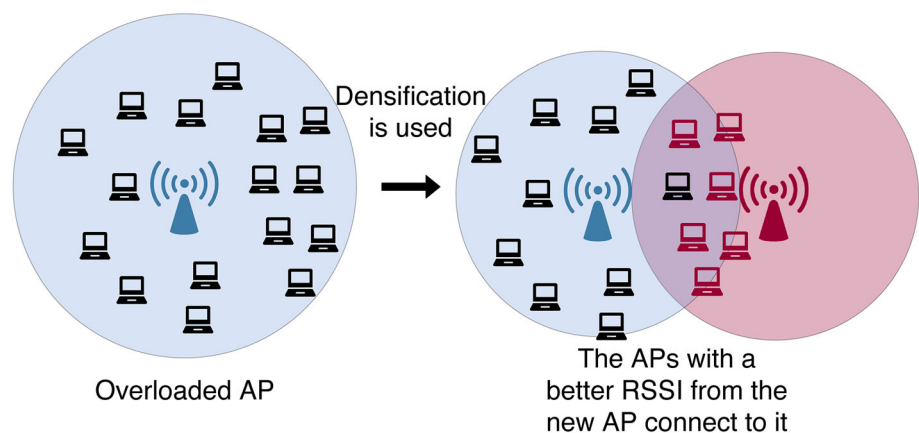
since SDN allows the proactive reaction of the network, it makes it possible to reassociate clients from overloaded APs to their underloaded neighbours in situations of uneven user distributions. By doing this, the stations make use of currently unused resources, eradicating the need to install new infrastructure. However, in the case of two APs in the same collision domain, looking at the load of the APs is not enough. In this case, the traffic handled by one of the APs sharing the collision domain affects that of the other. Thus, channel occupancy must be considered as well. Signal strength is also a key factor, as moving a client too far results in more retransmissions and a Modulation and Coding Scheme (MCS) with a lower physical rate that keeps the channel busy for a longer time, thus wasting resources.

In this work, we present Wi-Fi Multi-indicator Client Association (WiMCA), a user association scheme that improves the use of the available wireless resources to avoid the need for densification as a way to maintain performance. Taking the solution in [7] as a basis, the contribution of WiMCA consists in the avoidance of densification. For that purpose, the implementation in [7] has been changed to better estimate channel load considering the transmission rates and transmitted data in every time interval. It also aims to mitigate its effects when a big growth in the number of users all over the network makes densification inevitable. To do so, WiMCA jointly considers the three main factors that influence densification and the use of resources, namely AP load, RSSI and channel occupancy. The resulting solution bypasses the need for densification while avoiding the risk of being self-defeating. Ignoring any of the parameters could lead to an improvement in the performance in some situations but a deterioration in others. The performance evaluation in this work, which is designed to represent three scenarios that could lead to densification, shows that WiMCA is an improvement in network resource allocation over traditional CD approaches [8] and state-of-the-art approaches [9]. Moreover, the performance evaluation in [7] does not demonstrate all the strong points of the proposed solution such as the avoidance of congestion in the wireless resources, i.e., airtime. In this work, we carry out a more in-depth evaluation that allows analysing not only the issues found in [7] which has been improved by WiMCA and important aspects to be considered in dense networks such as AP load, mobility or channel occupancy.

The rest of the paper is organized as follows. Sect. 2 provides an overview of the related work. Sect. 3 describes WiMCA and its design principles. In Sect. 4, the results from the real-world performance evaluation are discussed. Finally, in Sect. 5, our conclusions are presented.

2 Related work

Association schemes in Wireless Local Area Networks (WLANs) have been the focus of much academic research, as illustrated by the existing body of literature. Furthermore, over the last few years, SDN has facilitated the inclusion of more efficient approaches without the need to modify the IEEE 802.11 standard. In this section we review the most relevant related work grouped by the strategy followed.

Bandwidth requirements. The works falling in this category take as main input the bandwidth requirements of the stations [10–14]. For instance, the authors of [10] propose a solution for load balancing across the APs that make up the network to achieve user bandwidth fairness. The problem is that each station must specify in advance its bandwidth requirements for the current session. The same target is pursued in [11], in which the authors propose an algorithm that aims to minimize the number of stations per AP based on the number of clients that each AP manages and the bandwidth needed by each one of them. However, the number of associated clients is not an accurate estimator of the workload of an AP. The above approaches may lead to ping-pong effects which are difficult to handle without coordination between the APs. This effect is tackled in [12], where a handover is not carried out until the same AP has been defined as the best choice for n consecutive times. The authors in [13] aim to maximize the bandwidth assigned to each client while keeping a balanced load. Bandwidth allocation is also used by the authors in [14]. In their approach, bandwidth demand differences between stations are taken into account. These approaches do not consider channel conditions and focus on improving a single parameter. They improve performance over the CD approaches, but they do not improve channel congestion or reduce collision domains so they do not free new resources.

Throughput maximization Other popular strategy pursued by some research works entails maximizing the aggregated throughput in the network [15–17]. Usually, stations tend to connect using a static distribution even if many different distributions are possible and may even offer a better performance. This is why the authors in [15] present a load balancing algorithm that selects different network interfaces for each flow on the basis of the required QoS to find the best possible path for all the flows. Some of the approaches using this strategy leverage SDN principles to improve their results [16, 17]. Graph theory is additionally used by the authors in [16] for throughput maximization by finding the best association graph between stations and APs. However, all the APs need to be in the same channel. The authors in [17] maximize

throughput by means of an abstraction of all virtual resources which they call Virtual Resource Chain. The throughput maximization strategy also focuses on a single parameter, but it does not reduce the amount of wasted resources due to interference or collisions.

Distributed approaches. Before the surge of SDN, distributed approaches were trendy because of the lack of a central controller providing a global view of the network and allowing its programmability. This lead some authors to opt for distributed approaches [18–20]. That is the case of [18], whose solution chooses the target AP on the station side by defining new metrics that take into account not only signal strength but also the performance degradation caused to other stations. In [19], authors propose an approach based on signal strength. Its decisions are taken on the station side. Likewise, a distributed algorithm is proposed in [20]. This approach measures load in deadline misses, in such a manner that the stations with a certain number of missed deadlines are reassociated to another AP. However, even though these distributed approaches outperform the standard CD association scheme, they are not optimal as they do not consider network-wide performance. Distributed approaches do not have a global view of the network and as a consequence their decisions might improve performance in a single AP while worsening performance on their neighbours.

Cell breathing Transmission power is another parameter that can be adjusted according to their load and their neighbours', resulting in the migration of stations from the most loaded APs to the least loaded ones. This method is known as Cell Breathing [21–23]. This is the approach that the work in [21] takes. The problem of this work resides on the difficulty of the implementation. An SDN-based centralized approach is presented in [22], which aims to improve the flexibility of Cell Breathing by taking advantage of the global view of the network. Another approach consisting in AP coverage adjustment is presented in [23].

Seamless-handovers Many strategies reassociate stations to redistribute load in the network. However, handovers produce a reassociation process that may involve a period when the station has no connection. Many works in the literature have proposed solutions to deal with this problem [24–28]. The authors in [24] propose a solution that requires that the APs operate on non-overlapping channels. The authors in [25] introduce the concept of Virtual Access Point (VAP) to allow seamless handovers. This allows them to set different channels for each AP as the clients have a dedicated AP that makes it possible to target beacons at a single client. In [26], the authors introduce the concept of Light Virtual Access Point (LVAP). LVAPs facilitate the state management of wireless stations. Each station has a unique LVAP. This allows seamless handovers between APs that operate on the same channel by replicating the LVAP at the destination AP and destroying the instance at the source AP. As the stations are not aware of the existence of the real AP, they remain connected to the same LVAP as if no change had been performed. LVAPs are also applied in [26–28].

Load balancing at AP-level Balancing load across the APs that form the network is another strategy found in the literature [29–37]. The authors of [29] use the average workload of the network to redistribute traffic when a new station joins the network. Nevertheless, modifications in the standard beacon frames are required. For this reason, most of the approaches using this strategy leverage SDN, which allows the programmability of the network and the use of LVAPs. The authors in [30] aim to balance the load at the AP level. Through an SDN centralized controller, they fill a scoring matrix that takes into account both the RSSI and occupancy rate. Similarly, the authors in [31] study the impact of a smart AP selection algorithm that uses a Fitingness Factor parameter. In [32], a reassociation algorithm is used to reduce interpacket delays and provide QoS on the basis of a Markovian analytical model. Likewise, an algorithm taking advantage of mobility and handovers is proposed in [33]. This approach defines an AP load threshold so that if an AP is above the threshold, it will not accept new stations. However, this approach limits user mobility. The approach in [34] forces handovers from the most loaded APs to the least loaded ones. In [35] the APs send load status reports to an SDN controller. This controller dispatches stations from overloaded APs to underloaded ones whenever a status change happens. However, just taking into account the load of each AP is far from optimal since many other factors affect channel conditions. Other approaches have tried to consider other KPIs to improve their decisions. Channel and AP congestion problems are aggravated in dense networks such as those of universities or office buildings, where Enterprise WLANs (EWLANs) are usually deployed. In this regard, a solution focused on EWLANs is presented in [36]. This approach takes into account the number of users per AP and the channel load at the AP level. This is inefficient as other APs in the same collision domain could overload the channel, resulting in throughput degradation for both APs. Even more metrics are combined in [37], where the RSSI, potential capacity, achievable data rate and the location of users are considered for association decisions. Just taking into account the load of each AP is far from optimal since many other factors affect channel conditions. APs are not isolated in dense networks and they may share collision domains with other APs. As a consequence, the traffic on each AP affects the other and thus, channel resources (airtime) become the bottleneck.

Up to this point, it seems clear that the amount of research carried out on this topic is vast. The approaches

relying on client reassociation to achieve load balancing at AP level. This is effective if there were no collision domains. However, the limited number of channels available in the IEEE 802.11 standard makes this impossible in very dense deployments. As a consequence, not only AP load but also channel load must be considered. If two APs share the same collision domain the air time used by one of them affects the other and vice versa. As a reaction to this, some approaches such as [38] present ways of minimizing collision domains. Other approaches such as [39] and [9] combine collision domain avoidance and smart user-association. However, they have shortcomings. In the case of [39], stations need to be modified. In the case of [9] it assigns channels to the APs that form the network to avoid collision domains, and then balances the load as if there were no collision domains. However, note that collision domains might be unavoidable, and as a consequence this might not be optimal.

In this paper, we present WiMCA, which solves these problems by achieving a trade-off between the load of the APs, the load on the channels and the signal quality keeping compatibility with off-the-shelf stations. Even though WiMCA minimizes collision domains, it takes into account the airtime in the channel instead of the airtime used by a single AP to avoid channel congestion problems in the cases where collision domains are unavoidable. Moreover, by considering the RSSI levels, it avoids the degradation of the average physical rate. This makes WiMCA more versatile and allows it to perform better in a wider range of situations. WiMCA is designed to be applicable in the real world as a solution for densification as it is able to fit more users and traffic in the same amount of resources, delaying the moment when densification is needed to cope with a general and massive increment in the number of users. More specifically, this work outperforms the state of the art in the following manners:

- WiMCA reduces collision domains and combines three indicators to be more versatile than the works focused on bandwidth requirements and maximizing throughput which focus on a single parameter. WiMCA not only makes better use of the available resources but also unlocks resources that were wasted due to interference or collisions.
- WiMCA leverages SDN principles to be able to program the network without the need to introduce modifications in the standard. This gives WiMCA a global vision of the network allowing it to take better reassociation decisions than distributed approaches.
- WiMCA keeps a constant transmission power and takes the most of the existent resources to avoid leaving stations out in the borders of the network as it happens in Cell Breathing approaches.

- Reassociations must be seamless to avoid performance drops. WiMCA applies seamless-handover strategy concepts like LVAPs to reassociate clients and achieve a better resource allocation.
- WiMCA tackles the problems of the load balancing at AP-level strategy by reducing collision domains, and taking into account the available airtime in the channel instead of the airtime of a single AP to avoid channel congestion when collision domains are unavoidable.

Finally, independently on the strategy followed, some approaches such as [15, 39] aim to improve versatility by using two interfaces or installing extra software on the stations. This reduces compatibility. One of the main goals of WiMCA is to achieve broad compatibility with the standard and off-the-shelf stations, so these kind of solutions are not considered in the design of WiMCA.

3 WiMCA: Wi-Fi multi-indicator client association

In this work we introduce Wi-Fi Multi-indicator Client Association (WiMCA), a new approach that defines three indicators aimed at monitoring resource allocation across the network in order to avoid densification, i.e., the need to install more APs to cope with the traffic demands while there are still free resources in the neighboring areas of the network, with the goal of studying new user distributions that make better use of the resources. For this purpose, the SDN-controller collects two major categories of information, namely the uplink RSSI and channel usage for each AP. By using these metrics, three indicators are defined to handle: (i) the load of the APs; (ii) the average RSSI of an AP; and (iii) the channel occupancy. By doing this, WiMCA reassociates clients to obtain a user distribution that avoids the need for more infrastructure by taking advantage of most of the resources already installed, thus avoiding the problems that densification implies.

Algorithm 1 User association algorithm

Input:

`AvRSSIs`: list of the average RSSI of every AP.

`APLoads`: list of the channel occupancy at every AP.

`ChOccup`: list of the channel occupancy at every channel.

Output:

`NewAP`: target AP to which a client is migrated.

1: **if** Max(`APLoads`)-Min(`APLoads`)>Med(`APLoads`) **then**

2: **for each** `i` in `APs` **do**

3:   **for each** `j` in `stationsAP`$_i$ **do**

4:    `newAP` ← Max$_j$(`RSSI`$_j$ × (`ChLoad`$_{AP_i}$ + `LoadAP`$_i$))

5: **else if** Max(`AvRSSIs`)-Min(`AvRSSIs`)>Med(`AvRSSIs`) **then**

6: **for each** `i` in `APs` **do**

7:   **for each** `j` in `stationsAP`$_i$ **do**

8:    `newAP` ← Max$_j$(`RSSI`$_j$ × (`ChLoad`$_{AP_i}$ +`LoadAP`$_i$))

9: **else if** Max(`ChOccup`)-Min(`ChOccup`)>Med(`ChOccup`) **then**

10:   ChannelReassignation()

Algorithm 1 represents a simplification of the WiMCA algorithm. Each one of the aforementioned indicators has an associated threshold that corresponds to the if statements in the algorithm. Every indicator triggers an action whenever the condition `Maximum - Minimum > Median` is fulfilled. This condition is checked periodically when new values are sent from an AP to the controller. This period is configurable on the controller side. Another important aspect is that the statistical functions, `Maximum`, `Minimum` and `Median` are not computed with the immediate values reported by the APs. Instead, the average of the history of the last 10 reported values is used. This is done to avoid ping-pong effects, as changes in a single report do not have an impact that is big enough to induce significant changes in the average of the historical series. The following subsections assess the different indicators in depth.

The first and the second indicators trigger a reassociation process that selects the best trade-off between signal quality and the network resources used. For this purpose, it reassociates clients to stations maximizing the value of `RSSI`$_j$× (`ChLoad`$_{AP_i}$+ `LoadAP`$_i$), where i refers to the AP and j refers to the station. We remind the reader that the value of RSSI is a negative number, and this is why the

product is maximized and not minimized. Good signal quality helps the selection of a higher MCS, which results in a reduction in the time taken for each frame to be transmitted, and as a consequence, the channel is free for longer and more frames can be transmitted via the channel. On the other hand, having more loaded APs and channels also has a negative impact on the value of the product. Therefore, the algorithm seeks the best combination of channel load and AP load. If there is no better option, there will be no migration.

This problem can be reduced to the Generalised Assignment Problem (GAP) [40]. In this problem there is a set of n tasks and m agents. Each task can be assigned to precisely one agent but an agent may be assigned to multiple tasks. The cost of the task depends on the agent-task assignment. The goal is to find the set of agent-task assignations that maximises profit. Similarly, in WiMCA, APs would be agents and stations would be tasks. WiMCA looks for the set of AP-station assignment that maximises the value of `RSSI`$_j$× (`ChLoad`$_{AP_i}$+ `LoadAP`$_i$). Every station can be assigned to precisely one AP but an AP may be assigned to multiple clients. Since GAP is NP-hard, WiMCA is NP-hard as well.

3.1 Indicator 1: AP load

The purpose of this indicator is to avoid overloaded APs. The usual solution in this situation is to install another AP in the area, making the cells smaller so that the clients, which usually follow a signal strength driven association scheme, are distributed over both APs. However, this requires an investment in new infrastructure and causes problems in the long term due to the small number of channels available in the IEEE 802.11 standard, especially in the 2.4GHz band, which could lead to APs sharing the same collision domain. If this is the case, the traffic of one AP affects the other, so congestion becomes unavoidable.

Finding a solution to avoid the need for installing more APs is very important, as the clients that are closer to a neighboring AP could be migrated to that AP even if the signal quality is lower provided that it is still sufficient for the communication to take place. CD approaches will never do that as they are not aware of resource usage. Note that, all clients benefit from such a migration, even those that have a worse signal quality than before, as they are connected to a less congested AP. As a consequence they can obtain more air time, sending more frames and enjoying a better performance. For the clients that maintain the associated AP, the fact that some clients are migrated to other APs reduces the congestion as the air time used by those clients at the source AP is now free, thus improving overall performance.

The AP load is measured using the% of channel occupancy of this AP which is approximated using Eq. 1 in contrast with [7] where the AP load was measured in bytes. This is not fully representative of the use of the wireless resources, since datarate and signal quality have an important influence on the use of the resources. For that reason channel occupancy is estimated using the total number of bytes that can be transmitted and the datarate of the stations. To do this, WiMCA uses the delivery probabilities provided by the Minstrel MCS selection algorithm [41] of each client. Since these probabilities together with the transmitted bytes (B) are periodically reported to the controller with a configurable period t, the estimated channel occupancy (O) of a client can be calculated using the rate with the best probability (R_{best}), as shown in Eq. 1. Note that this is a rough estimation which does not take into account many factors such as interframe spaces or contention windows, but it allows the occupancies of the different channels to be compared in a simple manner.

$$O(\%) = \frac{B \times 100}{R_{best} \times t} \tag{1}$$

Note that if this indicator reaches the aforementioned trigger threshold, a reassociation is studied and the other two indicators are not evaluated until the next report from the APs is sent to the controller, which is when Algorithm 1 will be executed again.

3.2 Indicator 2: average RSSI of an AP

The average RSSI of an AP refers to the average of the uplink RSSIs of all the clients connected to the AP. Thus, when this metric is low the physical rate provided by the MCS chosen by the MCS selection algorithm, such as Minstrel, will be lower in order to provide a more robust signal as a consequence of poor signal strength. Therefore, transmitted frames keep the medium busy for a long time. This wastes air time and does not help congestion. Moreover, upon the redistribution of clients carried out due to the AP load indicator, the average RSSI could become very low at some APs. If this is the case, more frames will need to be retransmitted.

To prevent this from happening, a new indicator to monitor this problem is added to study possible reassociations when this situation arises. The process carried out to study possible reassociations is the same as the one used in Indicator 1. Thus, when the average RSSI of an AP is above the threshold, WiMCA seeks a better distribution. In addition, WiMCA also checks possible redistributions upon individual worsening of the RSSI of a client. This is carried out to improve the performance in the case of moving stations, as new association possibilities that improve the performance may arise as the client moves.

3.3 Indicator 3: channel occupancy

WiMCA needs to rely on a channel assignment scheme that minimizes the number of APs in the same collision domain. As a consequence, an adaptive channel assignation scheme has to be defined. For this purpose, Channel Occupancy is used as a third indicator. Since WiMCA aims not only to avoid densification but also to mitigate its effect when a massive increment of users over the whole network makes it inevitable, this indicator aims to improve the performance of the network when there is more than one AP in the same collision domain, which usually happens in very dense deployments. This is done by avoiding the presence of overloaded channels.

An important change with respect to [7] is that channel occupancy is measured as the addition of the occupancy caused by the APs in the same collision domain, i.e., in carrier sense of each other and sharing the same channel. On the contrary, [7] only took into account the number of bytes being transmitted through the channel. In order to apply our solution to dense deployments where collision domains are probable, we need other solution because bytes do not offer a real vision of channel conditions, which are influenced by signal quality and Minstrel rate selection algorithm.

When this indicator is above the threshold, a channel reassignation is carried out using Algorithm 2. This channel reassignation assigns a channel to the AP with the lowest number of available channels in a recursive way. The available channels are those that have not been used yet by any of the neighbours of the AP. If any channel is available, the algorithm chooses the least used one to reduce interference and overlapping. However, if a channel is above the threshold, WiMCA removes the channel from the list of available channels, i.e., when the addition of the channel load generated by the APs using this channel goes above the threshold. This is necessary as not taking into account channel load may result in another channel distribution where one of the channels is still overloaded. However, with this modification, collision domains may not be minimal. This is why this algorithm is executed only when the situation cannot be solved through the other two indicators. Nevertheless, having more than one AP with a low load in the same collision domain is preferable to having an overloaded channel.

Algorithm 2 Channel assignment algorithm

Input:

`neighbours`: graph storing the neighbours of each AP.

`channels`: list of available channels.

`overlaps`: dictionary storing the overlapping channels.

`ChLoads`: list of each channel load.

Output:

`assignment`: tuples (AP, channel).

```
1: procedure CHANNELREASSIGNATION()
2:   remainingAPs ← APs not in assignment
3:   if Length(remainingAPs) == 0 then
4:     return assignment
5:   Sort remainingAPs by the lowest number of available
      channels and the highest of neighbors in assignment
6:   nextAP ← remainingAPs[0]
7:   possibleCh ← channels
8:   for each AP in neighbours[nextAP] do
9:     APCh ← assignment[AP]
10:    possibleCh ← possibleCh - APCh
11:    possibleCh ← possibleCh - overlaps[APCh]
12:   if not possibleCh then
13:    possibleCh ← Min(assignment)
14:   for each channel in possibleCh do
15:    assignment[nextAP] ← channel
16:    if Max(ChLoads) - Min(ChLoads) > Med(ChLoads) then
17:     possibleCh ← possibleCh - channel
18:    return  ChannelReassignation(neighbours, channel,
      assignment)
```

Algorithm 1 requires that each station in an AP goes through all the possible APs to find the best option. This has a complexity of $O(n \times m)$ where n is the number of APs and m is the number of stations. This is repeated for every AP, but to avoid incrementing complexity, the list of APs is ordered according to the value of $\text{RSSI}_j \times (\text{ChLoad}_{AP_i} + \text{LoadAP}_i)$. This algorithm makes the

recursive call to Algorithm 2. As a consequence, overall complexity is $O(n^2)$.

Algorithm 2 is a recursive procedure that ends when all APs have a channel assigned. First, on each recursion level, the remaining APs are sorted attending to the one that has the lowest number of available channels and the highest number of neighbors. Considering n as the number of APs, then the complexity of this is $O(n \cdot log(n))$. After that, the whole list of APs is traversed to discover available channels by removing the overlaping ones. The complexity of this part is $O(n)$. After that, the algorithm iterates through the list of possible channels which is no longer than n so the complexity is also $O(n)$. As a consequence, each recursive call takes $O(n)$. However, this is the time that each inner call takes. This function is called recursively $n - 1$ times. From that it can be inferred that, $T(n) = T(n-1) + O(n)$. Therefore, the complexity of this algorithm is $O(n^2)$.

4 Performance evaluation

4.1 Reference architecture

The implementation of the proposed scheme has been carried out taking as a reference the 5G-EmPOWER [5, 42] platform. 5G-EmPOWER combines SDN and Network Function Virtualization (NFV) in a single platform that supports lightweight virtualization and heterogeneous RATs such as Wi-Fi and LTE. This work is focused on Wi-Fi networks, in particular, EWLANs. It provides full visibility of the network state and allows the deployment of network services and fast prototyping of novel services and applications. The solution presented in this paper is deployed as a network application. This gives it a privileged position sitting on top of the controller, which allows it to obtain real-time information about the network. Every AP in the network keeps its state information so that the network can keep working in its last known state even if the central controller is not available.

One of the main advantages of leveraging SDN is the ease with which the required network metric can be gathered and the capacity to influence the behaviour of the network as a reaction to changes in those metrics. Periodically, the APs send the controller information about the channel usage, the clients and the RSSI, among other data and the period at which they do so is configurable on the controller side. However, this information is not used as soon as it is reported. Instead, the last 10 reports are kept and an average is computed, thus avoiding ping-pong effects. For example, traffic might be bursty, which could trigger an unwanted handover that has to be undone later

on. By using the average of the last 10 reports, the effect of an increase in channel occupancy or AP load should be consistent over time, meaning that only an impact big enough to significantly change the average should trigger a reassociation. Furthermore, 5G-EmPOWER provides LVAPs, which facilitates the state management of wireless stations through a high-level interface. Each station has its own unique LVAP [26]. This allows seamless handovers between APs even when they operate on different channels. If this is the case, the Channel Switch Announcements (CSAs) defined in the IEEE 802.11 standard are used to notify the client that it should start transmitting on the new channel.

Since each LVAP uses a specific Basic Service Set Identifier (BSSID), beacons and CSA frames are delivered in unicast mode to each user, which makes it possible to maintain both authentication and association active upon a handover. Handovers are carried out by instantiating an LVAP at the target AP, with this LVAP remaining inactive until the station connects it. Then, the old instance of the LVAP at the source AP is removed. This allows seamless handovers, a key feature for the solution presented in this paper, as it maintains the performance and simplifies the infrastructure management after reassociation decisions. In this work, our approach is applied to uplink transmissions. However, it could be applied in the same way to downlink transmissions by using IEEE 802.11k [43], although not many devices support this standard, except for some Apple ones [44]. This study is left for future work.

4.2 Evaluation methodology

The performance evaluation carried out in this work aims to demonstrate the effectiveness of the three indicators defined in WiMCA in better assigning resources in order to delay densification. To this end, the evaluation is composed of three experiments, one focused on each of the indicators introduced in Sect. 3. These experiments involve a deployment that may lead to densification. The methodology adopted is intended to reduce the variability obtained when analyzing all the indicators in the same experiment, which is already big enough in wireless networks, and isolate the effects of each indicator so that we can draw more accurate conclusions for each one. The evaluation was performed on a real testbed with three APs built on PCEngines ALIX 2D boards equipped with a single Atheros AR9220 Wireless Network Interface Controller (NIC). The experiments were carried out in the 2.4 GHz band using IEEE 802.11n. The APs were flashed with version 17.01 of EmPOWER-LEDE, an open-source Linux distribution based on OpenWRT [45].

The 5G-EmPOWER controller ran on a laptop with a quadcore Intel i5 processor running Linux Mint 19.2 Tina,

which was connected to the APs in a star topology. In the case of the clients, 15 Raspberry Pi 4's with 1 GB of RAM and Raspbian Buster were used. However, due to the high number of retransmissions caused by the integrated Wi-Fi interface of the Raspberry, a TP-Link TL-WN722N with a 4 dBi omnidirectional antenna was used instead. Two more clients were used as moving stations. In this case, they were Toshiba laptops powered by an Intel Atom processor and 4 GB of RAM running Ubuntu 19.04. Three additional laptops were set up in monitor mode to store the headers of all the frames on a channel, one for each channel. From these headers the key performance indicators were computed.

The area where the experiments were carried out was chosen to avoid external interference from other networks. The location of the APs was chosen to keep overlapping zones to allow seamless roaming between their coverage areas. The clients were placed leaving two within the range of each AP, and the rest at a middle point within the range of all the APs to see how the reassociations are performed. Nevertheless, each experiment presents certain inherent characteristics, such as the addition of 5 more stations, which will be explained in the following subsections. Table 1 summarizes the main data for each experiment, such as the channels used, the number of stations generating traffic in each part of the experiment (for those consisting of two parts), the number of moving stations, or the duration of the experiment.

In all the experiments, every client transmits at 3 Mbps to the server, which is located in the controller. That makes a total of 30 Mbps being transmitted over the network. This is an amount sufficient to congest one AP, but not enough to congest the whole network. As a consequence, a good resource allocation can avoid the congestion of one AP, which is the ultimate goal of WiMCA. If the network were completely congested, no changes in the performance would be detected, since all three APs would be congested and user and load redistributions would have no effect. We compare WiMCA with two state-of-the-art approaches: (i) a CD approach [8], and (ii) Wi-Balance [9]. This latter approach was chosen due to its good performance at load balancing at the AP level and code availability, as for many works the code is not available or does not provide enough information to reproduce the approach.

The goal of Experiment 1 was to generate a typical situation where densification is used to unload the overloaded AP. WiMCA should redistribute the load over the other APs when possible. In this experiment, we aim to see the effects of the AP load indicator. To show this behaviour, channels 1, 6 and 11 were used, to avoid collision domains so that the traffic of each AP does not affect the others. There are no moving stations in this experiment as we aim to avoid the variability that mobility would

Table 1 Test cases used in the performance evaluation

ID	Channels	Clients Part 1	Clients Part 2	Moving stations	Time
1	1, 6 and 11	10	15	–	60 s
2	1, 6 and 11	10	–	2	90 s
3	1 and 11	10	–	–	60 s

introduce in the experiment, making it more difficult to understand the results. The experiment consists of two parts, with each one taking 60 s. In Part 1, only 10 clients located at the positions depicted in Fig. 2 transmit. In Part 2, 5 more clients start transmitting. These clients are located at the coverage area of one of the APs so that the approaches applying load balancing move load to the neighboring APs.

Experiment 2 is intended to test how WiMCA reacts when there are changes in the signal quality of individual stations. To do so, two moving stations are used. These stations move on the path depicted in Fig. 3. The path was chosen so that the stations move within the coverage areas of all three APs, with the aim of forcing migrations due to low signal strength and testing the effect of the signal strength indicator on performance. The dots on the path are stops. These stations move for 5 s to the first stop and wait 10 s there. This process is repeated until the stations reach the end of the path. Channels 1, 6 and 11 are used. The duration of the experiment is 90 s as this is the time required to complete the path. Ten clients are constantly transmitting traffic to the AP and are located at the same place as in the first part of the previous experiment.

Finally, Experiment 3 is designed to test the channel occupancy indicator. Since WiMCA aims not only to avoid densification but also to mitigate its effect when a massive increment of users over the whole network makes it

unavoidable, this experiment aims to represent a situation where densification has already happened, causing two APs to share the same collision domain. The capacity to solve congestion in this kind of situations is an important feature of our algorithm. However, the performance evaluation in [7] does not demonstrate its benefits when there are APs in the same collision domain. As mentioned, this is solved by WiMCA by taking into account channel load instead of AP load. In this experiment, only channels 1 and 11 are used in order to force two of the APs to share the same collision domain. WiMCA is able to solve this situation as it looks for a trade-off between channel occupancy, signal strength and AP load. If it does not solve the situation by reassociation, WiMCA reassigns channels. Ten clients were located as shown in Fig. 4 in order to obtain an even distribution over the APs, as we wished to test the load distribution over the channels. The test takes 60 s with no moving stations and constant transmission for all the clients.

4.3 Results

In this subsection the results of the experiments are assessed. For each experiment, a set of key performance indicators are gathered. The first indicator is the aggregated goodput, i.e., the amount of useful information at the application level. Jain's fairness index is also computed to
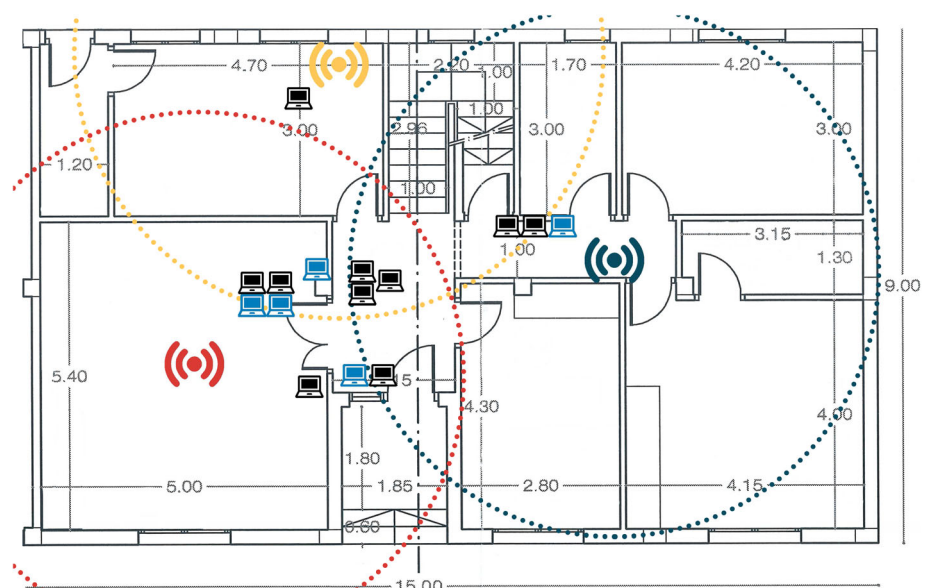

Fig. 2 Testbed layout for Experiment 1

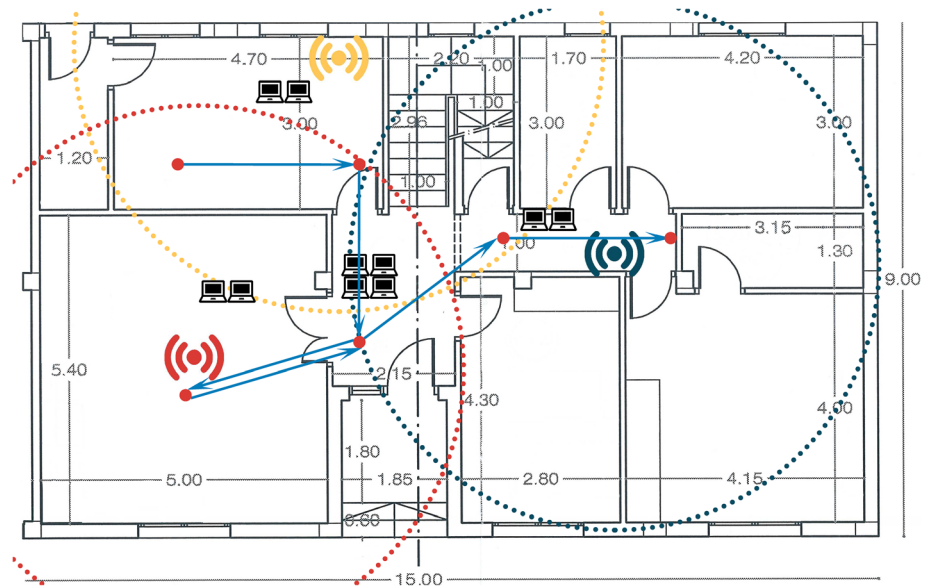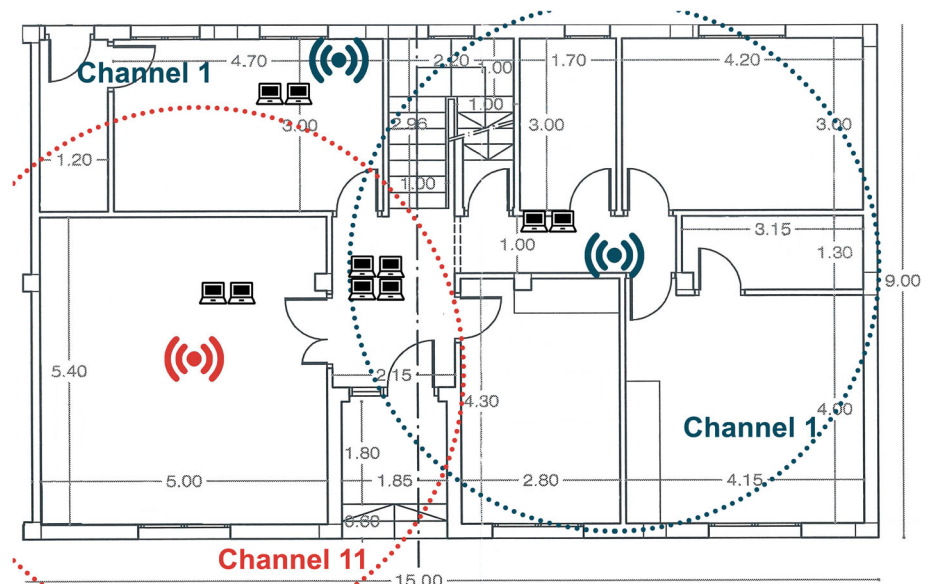**Fig. 3** Testbed layout for Experiment 2

Fig. 4 Testbed layout for Experiment 3

determine how fair the resource assignation carried out by the tested approaches is. Another important metric is the average number of retransmission attempts per transmitted frame, as this is indicative of both the number of collisions and the signal quality. The channel occupancy, i.e., the amount of channel time used in relation to the total time, is also important, but it needs to be assessed in relation to the average standard deviation of the channel occupancy, which refers to the deviation of the occupancy between the channels used in the experiment. Finally, the average physical rate selected by the MCS selection algorithm is assessed, as this is also indicative of the signal quality and the number of collisions in the network.

4.3.1 Experiment 1

The results of Experiment 1 are shown in Fig. 5. They show how WiMCA improves upon the other two approaches in both phases. This improvement is bigger in the second phase as the network load increases. In the case of the aggregated goodput shown in Fig. 5a, we can see that the better the load distribution, the greater the improvement in the aggregated goodput is, especially in Part 2 of the experiment, as there is less congestion and more traffic can be handled. This would allow the network administrator to delay the use of densification and to save costs when using WiMCA. WiMCA manages to improve aggregated goodput by 7% with respect to Wi-Balance and 50% with
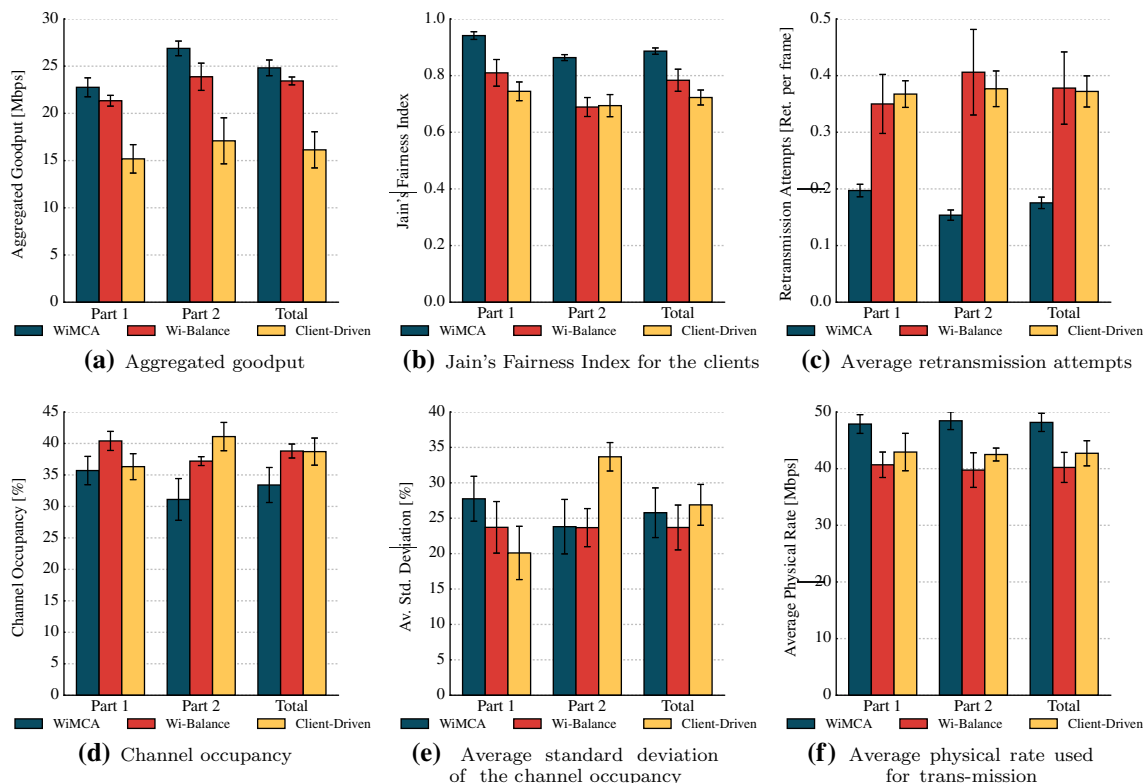
Fig. 5 Performance results of Experiment 1

respect to the CD approach in Part 1. In Part 2, this difference is 12% over Wi-Balance and 57% over the CD approach. This better resource assignation is visible also in Jain's Fairness Index, which is shown in Fig. 5b. In Part 1, WiMCA achieves 16% better fairness than Wi-Balance and 26% better than the CD approach. In Part 2, where the number of clients grows, this effect is even bigger as Wi-Balance obtains a similar performance to the CD approach and WiMCA improves upon the fairness of these two approaches by 25%.

Figure 5c shows the average number of retransmission attempts for each frame. The contention-based medium access of Wi-Fi leads to more retransmissions due to collisions when the channel occupancy is higher. Signal quality is also an important factor as fading and shadowing of the signal can also produce retransmissions. By taking into account these two factors, the trade-off achieved by WiMCA is effective, as it manages to reduce the number of retransmissions by half. This is in line with the reduction in the channel occupancy shown in Fig. 5d and the greater average physical rate shown in Fig. 5f. Wi-Balance manages to achieve a more similar use of the channels, but it sacrifices the signal quality for this as is shown by the reduction in its average physical rate, which WiMCA surpasses by 18% as the better signal quality leads Minstrel, i.e., the default rate control algorithm, to choose the

MCSs with higher rates and the greater number of retransmissions.

4.3.2 Experiment 2

Figure 6 presents the results of Experiment 2. They show that WiMCA adapts well to the movement of the stations, improving the results. On the conrary, the performance of Wi-Balance is negatively affected by the movement of the stations and the variation in signal quality, as shown in Fig. 6a, as it does not study the reassociation of clients to other APs when this factor deteriorates. As for the CD approach, its performance is affected by a less efficient resource allocation. This leads to congestion and a lower aggregated goodput. WiMCA achieves a good trade-off between these two factors, leading to 20% a better aggregated goodput. This better assignation of resources results in a higher Jain's Fairness Index, which allows all the clients to obtain the benefits of this augmented goodput, as shown in Fig. 6b.

Retransmissions, which are shown in Fig. 6c, are affected both by signal quality and collisions. We can see that Wi-Balance has a higher number of retransmission attempts than WiMCA and the CD approach as it does not use the RSSI to migrate clients. When looking at channel
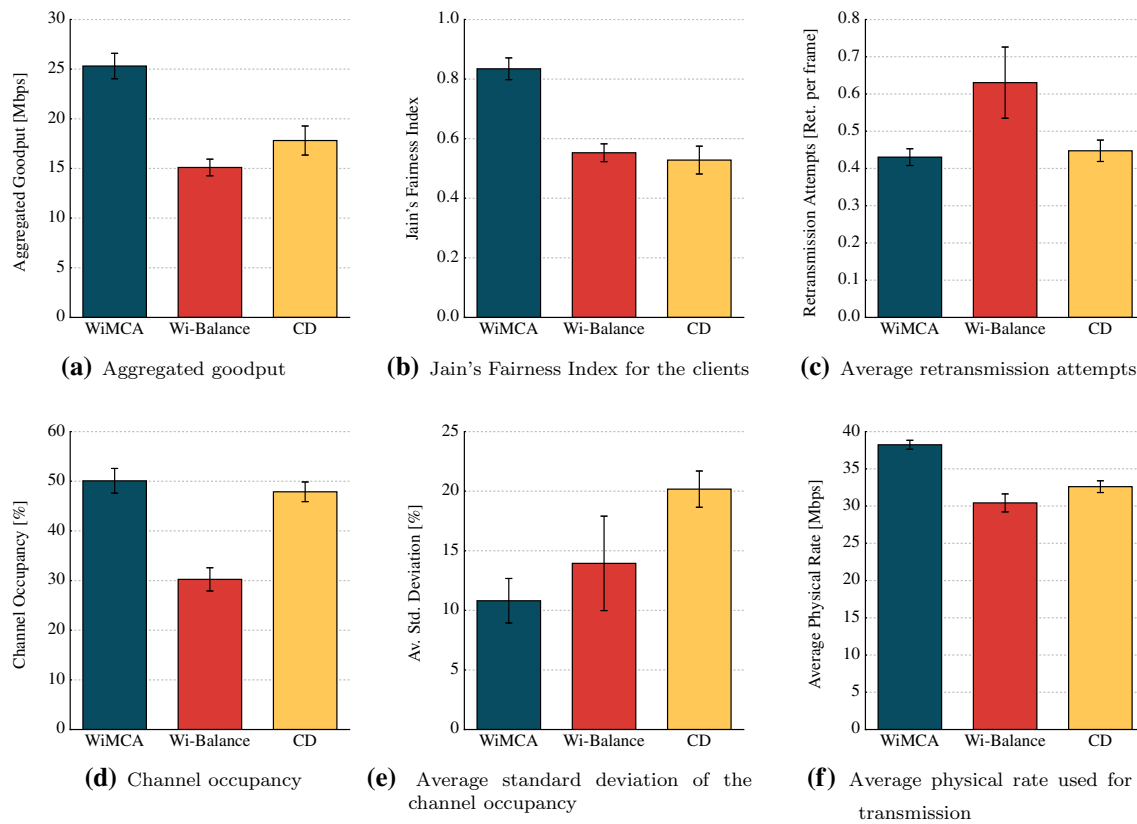
(a) Aggregated goodput

(b) Jain's Fairness Index for the clients

(c) Average retransmission attempts

(d) Channel occupancy

(e) Average standard deviation of the channel occupancy

(f) Average physical rate used for transmission

Fig. 6 Performance results of Experiment 2

occupancy, which is shown in Fig. 6d, Wi-Balance obtains a lower level of occupancy because in many cases it migrates clients to an AP that offers very bad signal quality, so they end up disconnecting. Figure 6e shows how WiMCA makes a more balanced use of the channels, surpassing Wi-Balance by 29% and the CD approach by 87%. The figure suggests that the CD approach is saturating some channels while leaving others underused. However, the worse distribution of the clients together with the lower MCS selected by Minstrel as a reaction to the collisions produced by this congestion is hindered by the better performance of the moving stations. However, WiMCA achieves the same occupancy. This is confirmed when looking at Fig. 6f, which shows that Minstrel selects an MCS with a 25% higher physical rate for WiMCA over Wi-Balance and 17% over the CD approach, both of which are affected by the aforementioned problems.

4.3.3 Experiment 3

Figure 7 presents the results of Experiment 3. It shows how WiMCA performs well by means of a good load balancing at the channel level. In particular, Fig. 7a shows that the aggregated goodput achieved by WiMCA is 16%

better than Wi-Balance and 37% better than the CD approach. Part of this increase in the aggregated goodput is thanks to a better use of the resources, as shown by Fig. 7d, where it can be observed that WiMCA uses channel 11 more efficiently. Channel 1 is saturated, since 2 APs are making use of it, so using the free airtime of channel 11 allows WiMCA to transmit more. This can be seen in Fig. 7e, which shows the standard deviation of the channel usage. WiMCA makes a more balanced use of the two channels, whereas Wi-Balance, which performs better than the CD approach, is not so efficient when dealing with channel occupancy. This also allows WiMCA to provide a higher Jain's Fairness Index, which gives the clients a fairer use of the resources so that they all can obtain the benefits of the augmented goodput, as shown in Fig. 7b. This happens because the clients that use channel 11 in Wi-Balance and the CD approach obtain more free airtime so they can transmit more than those on the saturated channel 1. WiMCA obtains a 10% higher fairness index than Wi-Balance, and a 26% higher fairness index than the CD approach.

By reducing the congestion in channel 1, WiMCA reduces the number of retransmissions by 4% with respect to Wi-Balance and by 20% with respect to the CD approach, as shown in Fig. 7c. Signal quality is very
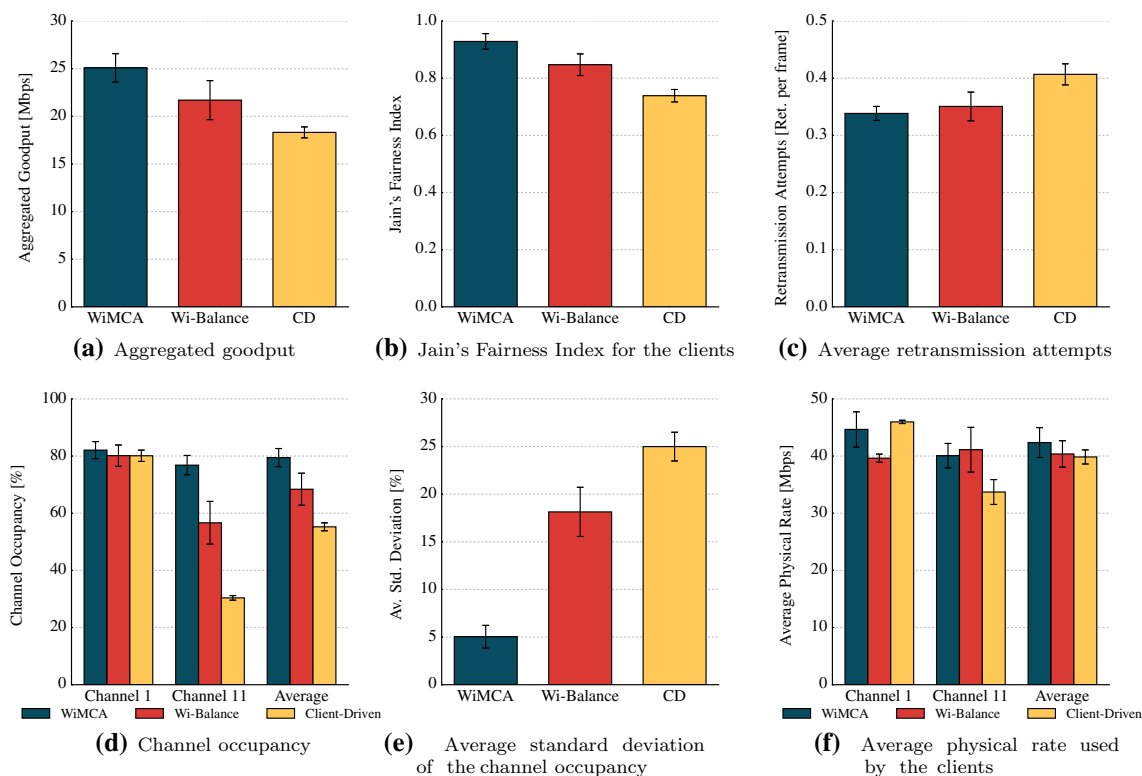
Fig. 7 Performance results of Experiment 3

important with regrards to the physical data rate used by the clients, but congestion also influences this, as Minstrel downgrades the MCS when there are retransmissions. Against this background, Fig. 7f shows how the trade-off is effective as even though WiMCA is not the best on any of the channels, it is the best on average. WiMCA obtains a higher average data rate, as it reduces collisions due to congestion and also keeps a good signal quality, whereas Wi-Balance, reduces collisions, but leaves signal quality unattended. The CD approach, for its part, does not take into account the load, so it does not avoid collisions due to congestion, but it does mantain a good signal quality.

4.3.4 Running time

In addition to the three previous experiments, the decision time has been measured. Decision time refers to the time since a report arrives to the controller from an AP until some decision is taken, i.e., migrate some station, redistribute channels, or taking no action. Time of 50 decisions was measured. The average decision time is 0.189 ± 0.088 s with 95% CI. The algorithm evaluates a possible redistribution every second. Thus, any decision time below 1 s is perfectly valid for WiMCA. Also, note that given that WiMCA works with user and load distributions, this is not strictly time constrained as those will not change fast.

5 Conclusions

In this paper, we have presented WiMCA, a client association scheme for Software-Defined Wi-Fi Networks that achieves a trade-off between the main factors affecting performance with the aim of avoiding the use of densification. The use of densification as a way maintain the performance when the number of users and the traffic increases is expensive and self-defeating as it makes collision domains unavoidable. If this happens, densification is not effective against congestion as the traffic of each AP affects the others. In order to delay the use of densification up to a general and massive increase in the number of users, WiMCA carries out a more efficient client association in order to make better use of the wireless resources. In particular, WiMCA achieves an effective trade-off between the average signal strength, the load of the APs and the load on the channels by defining one indicator for each one of these metrics.

The real-world performance evaluation shows that WiMCA improves the performance without the need for densification in several network configurations. WiMCA has been shown to outperform Wi-Balance, a state-of-the-art scheme, and the traditional CD approaches by increasing network aggregated goodput by up to 20% with respect to the former and up to 50% with respect to the latter. WiMCA also increases Jain's Fairness Index on the

client side, allowing most of the stations to obtain the benefits of this augmented goodput, showing that it is actually performing a better assignation of the resources. This better assignation is achieved by reducing retransmissions due to congestion and bad signal quality and by using those unused resources at underloaded APs. With this new allocation of resources more users and traffic are supported before reaching the point at which densification is needed, delaying its use to the worst case scenario. The results also show that WiMCA improves performance when densification has already been implemented and more than one AP is in the same collision domain.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Balachandran, A., Voelker, G. M., Bahl, P., & Rangan, P. V. (2002) Characterizing user behavior and network performance in a public wireless LAN. In *Proceedings of ACM SIGMETRICS performance evaluation review, Marina Del Rey, California* (Vol. 30, pp. 195–205).

2. Balazinska, M., & Castro, P. (2003). Characterizing mobility and network usage in a corporate wireless local-area network. In *Proceedings of ACM international conference on mobile systems, applications, and services (MobiSys), San Francisco, California, USA*.

3. Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., & Uhlig, S. (2014). Software-defined networking: A comprehensive survey. *Proceedings of the IEEE, 103*(1), 14–76.

4. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., et al. (2008). OpenFlow: Enabling innovation in campus networks. *SIGCOMM Computer Communication Review, 38*(2), 69–74.

5. SENSE, FBK . (2020). *The 5G-EmPOWER wiki.* https://github.com/5g-empower/5g-empower.github.io/wiki. Accessed on June 18, 2020.

6. Suresh, L., Schulz-Zander, J., Merz, R., Feldmann, A., & Vazao, T. (2012). Towards programmable enterprise WLANS with Odin. In *Proceedings of the first workshop on Hot topics in software defined networks, Association for Computing Machinery, New York, NY, USA, HotSDN '12* (pp. 115–120).

7. Gómez, B., Coronado, E., Villalón, J., Riggio, R., & Garrido, A. (2020). User association in software-defined wi-fi networks for enhanced resource allocation. In *Proceedings of IEEE Wireless communications and networking conference (WCNC), Seoul, South Korea*.

8. Gong, D., & Yang, Y. (2012). AP association in 802.11n WLANs with heterogeneous clients. In *Proceedings of IEEE international conference on computer communications (INFOCOM)*.

9. Coronado, E., Riggio, R., Villalón, J., & Garrido, A. (2018). Wi-balance: channel-aware user association in software-defined Wi-Fi networks. In *Proceedings of IEEE/IFIP network operations and management symposium (NOMS), Taipei, Taiwan*.

10. Sun, L., Wang, L., Qin, Z., Yuan, Z., & Chen, Y. (2018). A novel on-line association algorithm for supporting load balancing in multiple-AP wireless LAN. *Mobile Networks and Applications, 23,* 1–12.

11. Sheu, S. T., & Wu, C. C. (1999). Dynamic load balance algorithm (DLBA) for IEEE 802.11 wireless LAN. *Journal of Applied Science and Engineering, 2*(1), 45–52.

12. Yen, L. H., & Yeh, T. T. (2006). SNMP-based approach to load distribution in IEEE 802.11 networks. In *Proceedings of IEEE vehicular technology conference (VTC), Melbourne, Vic., Australia*.

13. Bejerano, Y., Han, S. J., & Li, L. E. (2004). Fairness and load balancing in wireless LANs using association control. In *Proceedings of international conference on mobile computing and networking, Philadelphia, PA, USA*.

14. Yang, L., Cui, Y., Tang, H., & Xiao, S. (2015). Demand-aware load balancing in wireless lans using association control. In *Proceedings of IEEE global communications conference, San Diego, CA, USA*.

15. De Schepper, T., Latré, S., & Famaey, J. (2017). A transparent load balancing algorithm for heterogeneous local area networks. In *Proceedings of IFIP/IEEE international symposium on integrated network management (IM), Lisbon, Portugal*.

16. Lei, T., Wen, X., Lu, Z., & Li, Y. (2017). A semi-matching based load balancing scheme for dense IEEE 802.11 WLANs. *IEEE Access, 5,* 15332–15339.

17. Feng, J., Zhao, L., Chen, C., Ren, Z., & Du, J. (2016). User-oriented load balance in software-defined campus WLANS. In *Proceedings of IEEE vehicular technology conference (VTC), Nanjing, China*.

18. Abusubaih, M., Gross, J., Wiethoelter, S., & Wolisz, A. (2006). On access point selection in IEEE 802.11 wireless local area networks. In *Proceedings of IEEE conference on local computer networks (LCN), Tampa, FL, USA*.

19. Papaoulakis, N., Patrikakis, C., Stefanoudaki, C., Sipsas, P., & Voulodimos, A. (2011). Load balancing through terminal based dynamic AP reselection for QoS in IEEE 802.11 networks. In *Proceedings of IEEE conference on pervasive computing and communications (PerCom), Seattle, WA, USA*.

20. Collotta, M., Pau, G., Salerno, V.M., & Scatà, G. (2012). A distributed load balancing approach for industrial IEEE 802.11 wireless networks. In *Proceedings of IEEE international conference on emerging technologies and factory automation, Krakow, Poland*.

21. Garcia, E., Vidal, R., & Paradells, J. (2008). Cooperative load balancing in IEEE 802.11 networks with cell breathing. In *Proceedings of IEEE symposium on computers and communications (ISCC), Marrakech, Morocco*.

22. Lin, C. Y., Tsai, W. P., Tsai, M. H., & Cai, Y. Z. (2017). Adaptive load-balancing scheme through wireless SDN-based association control. In *Proceedings of IEEE international conference on advanced information networking and applications (AINA), Taipei, Taiwan*.

23. Wang, S., Huang, J., Cheng, X., & Chen, B. (2014). Coverage adjustment for load balancing with an AP service availability guarantee in WLANs. *Wireless Networks, 20*(3), 475–491.

24. Chan, Y. C., & Lin, D. J. (2014). The design of an ap-based handoff scheme for IEEE 802.11 WLANs. *International Journal of e-Education, e-Business, e-Management and e-Learning, 4*(1), 72.

25. Berezin, M. E., Rousseau, F., & Duda, A. (2011). Multichannel virtual access points for seamless handoffs in IEEE 802.11 wireless networks. In *Proceedings of IEEE vehicular technology conference (VTC), Yokohama, Japan*.

26. Suresh, L., Schulz-Zander, J., Merz, R., Feldmann, A., & Vazao, T. (2012). Towards programmable enterprise WLANS with Odin.

In *ACM hot topics in software defined networking (HotSDN)*, Helsinki, Finland.

27. Rangisetti, A. K., Baldaniya, H. B., Kumar, P., & Tamma, B. R. (2014). Load-aware hand-offs in software defined wireless LANs. In *Proceedings of IEEE international conference on wireless and mobile computing, networking and communications (WiMob)*, Larnaca, Cyprus.

28. Zeljković, E., Slamnik-Kriještorac, N., Latré, S., & Marquez-Barja, J. M. (2019). ABRAHAM: Machine learning backed proactive handover algorithm using SDN. *IEEE Transactions on Network and Service Management, 16*(4), 1522–1536.

29. Jabri, I., Krommenacker, N., Divoux, T., & Soudani, A. (2008). IEEE 802.11 Load balancing: an approach for QoS Enhancement. *International Journal of Wireless Information Networks, 15*(1), 16–30.

30. de Sa, F.R., da Cunha, A. M., & Cesar, C. d. A. C. (2017). Effective AP association in SDWN based on signal strength and occupancy rate. In *Proceedings of IEEE conference on local computer networks (LCN)*, Singapore.

31. Raschellà, A., Bouhafs, F., Mackay, M., Zachariou, K., Pilavakis, V., & Georgiades, M. (2018). Smart access point selection for dense WLANs: A use-case. In *Proceedings of IEEE wireless communications and networking conference (WCNC)*, Barcelona, Spain.

32. Chen, J., Liu, B., Zhou, H., Yu, Q., Gui, L., & Shen, X. (2017). QoS-driven efficient client association in high-density software-defined WLAN. *IEEE Transactions on Vehicular Technology, 66*(8), 7372–7383.

33. Kiran, N., Changchuan, Y., & Akram, Z. (2016). AP load balance based handover in software defined WiFi systems. In: *Proceedings of IEEE international conference on network infrastructure and digital content (IC-NIDC)*, Beijing, China.

34. Sood, K., Liu, S., Yu, S., & Xiang, Y. (2015). Dynamic access point association using software defined networking. In *Proceedings of international telecommunication networks and applications conference (ITNAC)*, Sydney, NSW, Australia

35. Lin, Y. D., Wang, C. C., Lu, Y. J., Lai, Y. C., & Yang, H. C. (2018). Two-tier dynamic load balancing in SDN-enabled Wi-Fi networks. *Wireless Networks, 24*(8), 2811–2823.

36. Han, Y., Yang, K., Lu, X., & Zhou, D. (2016). An adaptive load balancing application for software-defined enterprise WLANs. In *Proceedings of international conference on ICT convergence*, Jeju, South Korea.

37. Cao, F., Zhong, Z., Fan, Z., Sooriyabandara, M., Armour, S., & Ganesh, A. (2016). User association for load balancing with uneven user distribution in IEEE 802.11ax networks. In: *Proceedings of IEEE consumer communications & networking conference (CCNC)*, Las Vegas, NV, USA.

38. Jeunen, O., Bosch, P., Van Herwegen, M., Van Doorselaer, K., Godman, N., & Latré, S. (2018). A machine learning approach for IEEE 802.11 channel allocation. In: *Proceedings of IEEE international conference on network and service management (CNSM)*, Rome, Italy.

39. Oh, H. S., Jeong, D. G., & Jeon, W. S. (2020). Joint radio resource management of channel-assignment and user-association for load balancing in dense WLAN environment. *IEEE Access, 8,* 69615–69628.

40. Kundakcioglu, O. E., & Alizamir, S. (2009). *Encyclopedia of optimization, generalized assignment problem* (pp. 1153–1162). Berlin: Springer.

41. Xia, D., Hart, J., & Fu, Q. (2013). Evaluation of the Minstrel rate adaptation algorithm in IEEE 802.11g WLANs. In *IEEE international conference on communications (ICC)* (pp. 2223–2228). IEEE.

42. Riggio, R., Marina, M. K., Schulz-Zander, J., Kuklinski, S., & Rasheed, T. (2015). Programming abstractions for software-defined wireless networks. *IEEE Transactions on Network and Service Management, 12*(2), 146–162.

43. (2008) Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 1: Radio Resource Measurement of Wireless LANs. *IEEE Std 80211k-2008 (Amendment to IEEE Std 80211-2007)* (pp. 1–244).

44. Sanchez, M. I., & Boukerche, A. (2016). On IEEE 802.11 K/R/V amendments: Do they have a real impact? *IEEE Wireless Communications, 23*(1), 48–55.

45. Endt, T., Brown, R., & Bursi, A. (2020). *OpenWRT project*. https://openwrt.org/es/start. Accessed on June 17, 2020.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Blas Gómez is a PhD student in Advanced Computer Technologies at the Universidad de Castilla-La Mancha, Spain. He graduated in Computer Science in 2018 in the same university and received his MSc in Computer and Networks Engineering from the Politechnical University of Valencia in 2019. Also in 2019, he joined the High-Performance Networks and Architectures Group in the University of Castilla-La Mancha as pre-doctoral researcher. His current research interests and areas of publication include multimedia delivery over WLAN and SDN WLANs.

Estefanía Coronado is a Senior Researcher at i2CAT Foundation (Spain). From 2018 to 2020 she was an Expert Researcher at Fondazione Bruno Kessler (Italy). In 2018 she completed her PhD at the University of Castilla-La Mancha (Spain) on multimedia distribution over SD-WLANs using ML techniques. She received the M.Sc. degrees in Computer Engineering and Advanced Computer Technologies in 2014 and 2015 from the same university. She has published around 30 papers in international journals and conferences. Her research interests include AI-driven network automation, wireless/mobile networks, network slicing, MEC-enabled systems and SDN/NFV.

José M. Villalón received the Ph.D. degree in computer engineering from the University of Castilla–La Mancha, Spain, in 2007. In 2007, he joined the Department of Computer Systems, University of Castilla–La Mancha, where he is currently an Associate Professor. He has also been a Visiting Researcher with INRIA, Sophia Antipolis, France. His research interests include high-performance networks, wireless networks, quality of service and quality of experience over Wi-Fi and WiMAX, multicast transmission, software defined networking, video transmission, and error–resilient protocol architectures.

Roberto Riggio is Senior Researcher in the Connected Intelligence Group at RISE AB in Stockholm, Sweden. He got his PhD in distributed systems under the supervision of Prof. Imrich Chlamtac from the University of Trento (Italy). From 2008 to 2017 he held several positions at CREATE-NET in Trento (Italy) including the Chief 5G Scientist role. In 2018 and 2019 he was head of the Wireless and Networked Systems Unit at FBK and in 2020 he was Head of the Smart Networks and Services Unit also at FBK. Finally, in 2020 he was Senior 5G Researcher at the i2CAT Foundation in Barcelona, Spain. From 2018 to 2020, within the EU Horizon 2020 5G-CARMEN, he coordinated the first world-wide cross-country validation of 5G for connected, cooperative and automated mobility (CCAM) across the Bologna to Munich 5G Corridor. He has published more than 130 papers (h-index 25) in internationally refereed journals and conference proceedings and has generated more than 4.3 MEuro in competitive funding in the last 5 years. He has received several awards including the IEEE INFOCOM Best Demo Award (2013 and 2019) and the IEEE CNSM Best Paper Award (2015). He serves on the TPC/OC of leading conferences in the networking field and is an associate editor for the Wiley International Journal of Network Management, the Springer Wireless Networks journal, and the IEEE Transactions on Network and Service Management. He is a member of the ACM and a Senior Member of the IEEE.

Antonio Garrido is a Full Professor at the University of Castilla-La Mancha, Spain. He has led several research projects in telemedicine, computer networks, and advanced computer architectures. He collaborates with the Spanish Education Ministry on the quality evaluation of Spanish universities since 2000. He was the Dean of the School of Computer Engineering. His research interests include high-performance networks, Wireless networks and video compression. He has published over 50 papers in international journals.