Reinforcement Learning to Enhance Spatial Reuse in Dense Wireless Networks

A thesis proposal

by

Francesc Wilhelmi

supervised by Boris Bellalta, Cristina Cano & Anders Jonsson

Department of Information and Communication Technologies
Universitat Pompeu Fabra, Barcelona

September 2017

Contents

Chapter 1

INTRODUCTION

1.1 Motivation

Due to the growing popularity of wireless deployments, specially the ones based in the IEEE 802.11 standard, it is very common to find overlapping Wireless Networks (WNs) that squander the bandwidth resources. This is mostly provoked by the usual lack of organization and/or agreement when sharing resources such as the available spectrum. In addition, the performance issues get worse as the number of overlapping networks increases. According to [1], the performance in a WN is severely deteriorated as density increases, specially if two or more Overlapping Basic Service Sets (OBSSs) coexist. A clear evidence about the aforementioned neighbours disregarding can be found in the power level used by transmitters, which is, by default, typically set to the maximum, regardless of the distance between nodes and the channel occupancy [2]. If we consider dense scenarios such as stadiums, trains or apartment buildings, minimising coexistence issues turns out to be a challenging task. To overcome the coexistence issues in WNs, three domains are mainly being studied:

- **Time:** current IEEE 802.11 WLANs are governed by the Distributed Coordination Function (DCF), which grant temporal channel access through the Carrier Sense Multiple Access with Collision Avoidance (CSMA/ CA) protocol. However, this approach is well-known to generate performance issues that lead to starvation and/or collisions. A step further in the CSMA/CA protocol is found in CSMA with Enhanced Collision Avoidance (CSMA/ECA), which determines the backoff counter after successful transmissions. [3] shows the benefits in terms of performance when using CSMA/ECA instead of CSMA/CA. An interesting alternative to CSMA is presented in [4], which proposes Time Frequency CSMA/CA (TF-CSMA/CA), an algorithm that schedules packets in both time and frequency domains. Using TF-CSMA/CA, a station also performs backoff in the frequency domain, which is shown to increase the channel utilisation. However, it requires a high self-organization degree to find appropriate spectrum bands that allow minimising the interference. Another alternative to CSMA/CA is presented in [5], which aims to identify hidden and exposed nodes at MAC, and transmit control information at PHY.

- **Frequency:** due to the scarcity of the available spectrum, it is required to smart mechanisms to make the most of the frequency resources. This specially necessary in dense scenarios, so that channel utilisation can be maximised while minimising the total interference. For that purpose, we find Dynamic Channel Bonding (DCB) and Dynamic Spectrum Access (DSA). DCB is a promising technique for dynamically increase the width of the utilised spectrum during transmissions, which allows increasing the throughput experienced by a given WN. Furthermore, DSA is focused in finding reliable free spectrum patterns, which is useful to provide an optimal frequency allocation in overlapping WNs.

- **Space:** the spatial domain considers the transmission range at which a given transmitter is able to reach due to the strength of its transmission power, as well as the sensitivity area in which the receiver is able to detect incoming transmissions. If both transmission range and sensitivity area are within the same space, and both transmitter and receiver are in different WNs, the sensed power is considered as interference. Due to the DCF operation, overlapping transmissions lead to starvation or collisions, according to the situation. If a node suffering from interference is about to send a packet, it must defer its transmission in case of sensing the channel busy (exposed node problem). Similarly, if the node was also receiving data from another device, it can suffer a collision if the sensed interference is higher than its capture effect (hidden terminal problem). To enhance spatial reuse, we find techniques such as Transmit Power Control (TPC), Carrier Sense Threshold (CST) adjustment, and directional transmissions, which are further described in Section 2.3.

1.2 Open Challenges

To the matter of this thesis, we highlight the enhancement of spatial reuse, which is specially emphasised in the IEEE 802.11ax-2019 standard that is being developed by the Task Group ax (TGax). Spatial reuse aims to reduce the interference level in an overlapping scenario, so that the area throughput can be maximised. To enhance spatial reuse, we focus on TPC, CST adjustment, and single and multi-user beamforming. We next list the open challenges for spatial reuse improvement in future WNs:

- Bandwidth requirements: there is an increasing trend in new applications of relying in communications to deliver their content, since they increasingly rely in the computational power of the cloud.

- High-density environments: the irruption of new mobile devices results in new scenarios, so that coexistence presents unprecedented challenges. Henceforth, novel mechanisms and architectures are required to confront the over-population of mobile devices.

- Chaotic deployments: in addition to density, another open challenge consists in facing chaotic deployments in which several WNs compete for the same resources in an uncoordinated manner. Since the current decentralised solutions lack of effectiveness in dense environments, collaboration between independent WNs turns out to be necessary to solve the resource allocation problem.

- Underutilisation of the limited spectrum: typically, WLANs fail in taking full advantage of frequency resources (idle periods, collisions, suboptimal spatial reuse, etc.). The lack of self-adaptive mechanisms in wireless devices prevents capturing the changes in the environment, which leads to ineffective configurations.

- Network variability: not only users arrivals and departures entails a challenge in WNs, but the bandwidth requirements variability imply real-time solutions to accommodate the available resources.

To face the abovementioned open challenges, many techniques have been developed (summarised in Figure 1.1), from which we highlight the following ones, particularly focused on IEEE 802.11 WLANs:

- Contention Window (CW) Adaptation: in order increase efficiency at channel utilisation in terms of time, we find CW adaptation. This technique allows minimising the number of empty slots, thus increasing the throughput. However, a too small CW may lead to collisions, which may be counterproductive.

- Scheduled approaches: a step further for empty slots minimisation is found in scheduled approaches. In them, the access to the channel is managed through a central unit or a coordinated set of devices, which guarantees a high channel utilisation without experiencing collisions. Notwithstanding, scheduled approaches require from communication overhead, which may be counter productive.

- Dynamic Spectrum Access (DSA): due to the scarcity of the available spectrum, making a good channel allocation is necessary to improve the network performance in dense environments. Henceforth, to maximise the spectral reuse, it is necessary to assign channels in a such way that interference is minimised.

- Dynamic Channel Bonding (DCB): occupying more than one channel for a single transmission is also an interesting technique to maximise the throughput of a wireless network. However, it is not trivial to establish a channel bonding policy, since using more than one channel may harm the overall throughput in overlapping scenarios. In addition, the transmission power is reduced proportionally with the number of channels used, which also complicates the decision-making process.

- Transmission Power Control (TPC): increasing/decreasing the power transmitted is useful to reach larger distances or to reduce the generated interference, respectively. For the latter, it is worth to mention that spatial reuse can be enhanced if short communications use the only necessary transmission power. If so, interference is minimised and the number of parallel transmissions can be increased.

- Sensitivity adjustment: adjusting the sensitivity in a given device allows increasing the transmission range or reducing it, according to the situation. For the former, decreasing the sensitivity allows listening to farther devices. However, it also leads to a lower spatial reuse, thus minimising the number of parallel transmissions. Furthermore, increasing the sensitivity is useful to restrict the transmission range, leading to a higher number of short-range transmissions. Nevertheless, the collisions probability by hidden nodes may be increased.

- Directional transmissions: to even improve spatial reuse, directional transmissions reduce the interference area of a given transmission. The fact of directing the beam of energy towards the receiver allows reducing the number of exposed nodes, thus maximising the number of parallel transmissions. However, in IEEE 802.11 WLANs, the synchronism may be lost due to the fact that not all the devices listen to the ongoing transmission.

- Interference cancellation: a step further in spatial reuse maximisation is found in interference cancellation, which aims to extract an interference-free signal from a set of incoming transmissions. Interference cancellation allows multiple simultaneous transmissions.

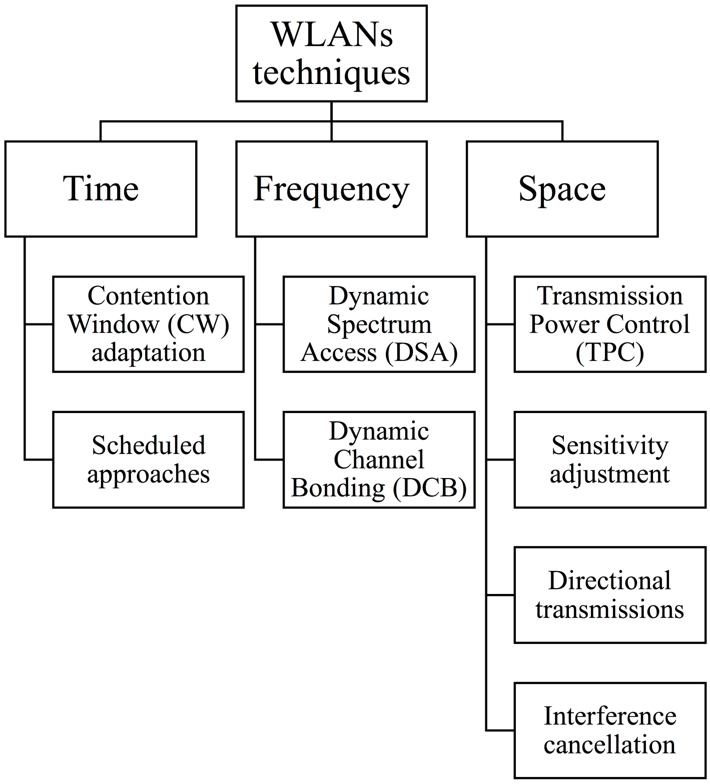


Figure 1.1: Techniques to enhance WLANs performance

1.3 Contributions

Given the open challenges and problems regarding current and future wireless deployments, the expected contributions of this thesis are:

- Understanding the challenges and limitations of future WLANs, thus allowing researchers to deeply understand the topic and provide new solutions. To do so, we aim to perform an exhaustive analysis of the state-of-the-art in spatial reuse, as well as in the usage of Artificial Intelligence (AI) in WNs.

- Exploring the feasibility of applying AI techniques to enhance spatial reuse in WNs, i.e., studying its practicality to show if they allow improving the performance of a given WN. In addition, we aim to show the benefits and limitations of applying a set of techniques for different architectures (e.g., fully centralised, uncoordinated, etc.). Regarding the network architecture, it is important to deeply analyse in which devices must a candidate solution be implemented, as well as in which of the protocols layer (e.g., PHY, MAC, Network, etc.). For the latter, the decision must be done according to the available input information in each level and the impact on the network performance.

- Designing decentralised ML-based solutions for enhancing spatial reuse, so that WNs become autonomous systems through self-configuration. Moreover, it is also of importance to provide real-time mechanisms that can satisfy the challenge that variability in WNs supposes. We aim to explore both coordinated (e.g., with message passing) and uncoordinated solutions, motivated by the chaotic scenarios in which several independent WNs coexist (e.g., residential buildings).
- As an extension of the previous point, we also aim to provide centralised ML-based solutions, so that many other fully controlled scenarios (e.g., university campus, airport, etc.) are covered. In such environments, it is important to analyse the trade-off between the optimality of the solution and the required time to achieve it.
- Integrating the abovementioned solutions with others, such as Dynamic Channel Bonding, to provide a higher improvement in WNs performance.

1.4 Document Structure

The remainder of this document is structured as follows: Section 2 presents the Research Problem, as well as open challenges regarding the spatial reuse problem in WNs. The related work is shown in Section 3, which also includes the current literature in the application of Reinforcement Learning (RL) to Wireless Communications. Then, Section 4 presents the ongoing work and the current contributions to the matter of this thesis. Finally, the working plan and the expected contributions are detailed in Section 5.

Chapter 2

RESEARCH PROBLEM

In this Section we describe the current issues in coexisting IEEE 802.11 WLANs, as well as the open challenges for the spatial reuse enhancement. Note as well, that IEEE 802.11 Wireless Local Area Networks (WLANs) are particularly focused to narrow the problem characterisation.

2.1 Coexistence Issues in IEEE 802.11 WLANs

Devices in IEEE 802.11 WLANs implement the Distributed Coordination Function (DCF) for accessing to the channel in presence of other nodes using the same frequency band. DCF combines Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) with Binary Exponential Backoff (BEB). Henceforth, before a packet transmission, a station must listen to the channel for a period called Distributed Inter Frame Space (DIFS). Channel is sensed to be free according to the Clear Channel Assessment (CCA) mechanism, i.e., if the power sensed is lower than a given threshold. The power received at a given node is the sum of all the interference generated by the other devices under the environment-constrained propagation effects. Regarding BEB, it aims to reduce the number of potential collisions by randomizing the access to the channel, and by adapting the opportunities for accessing to it. In BEB, each station selects a random backoff value to start a countdown that prevents to transmit until it has been exhausted, thus minimizing collisions. In case channel is sensed as busy, the countdown is paused. It is resumed as soon as the ongoing transmission finishes and the channel is sensed free again. Note that the computed backoff is a uniformly distributed random variable between 0 and the Contention Window (CW), which is doubled in case of noticing a packet failure until reaching a maximum value CW$_{max}$. The probability of noticing a collision decreases as CW increases, but entails higher idle periods (also leading into throughput degradation). The DCF operation is summarised in Figure 2.1

Despite the current mechanisms for allowing coexistence in IEEE 802.11 WLANs, and due to the nature of the CSMA/CA protocol, networks overlapping still drives into many problems and situations that result into poor throughput performance. In particular, there is a trade-off between the collisions occurring in an OBSS, the time channel remains idle, and the overhead generated to avoid harmful situations. In all the cases, both throughput and fairness in a WLAN are compromised. Two well-known problems often occur in IEEE 802.11 coexisting deployments, which are a matter of interest for the research community. The first one, the Exposed-Terminal problem (or exposed node problem), occurs when a device is prevented to

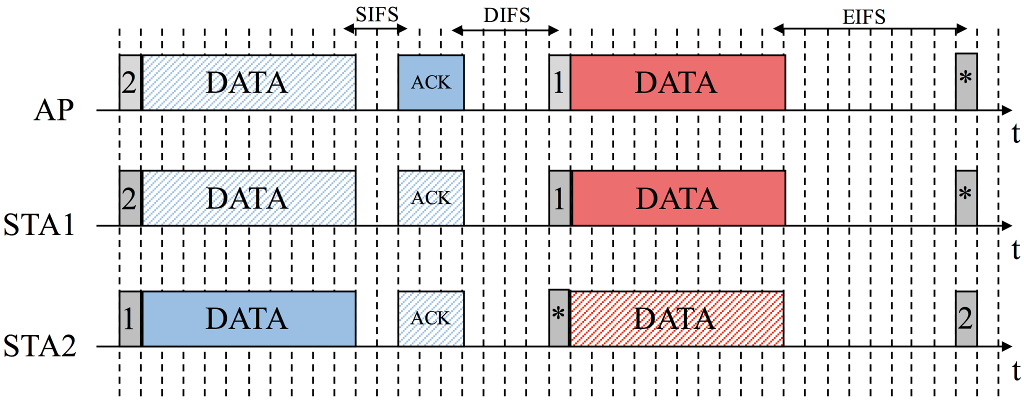


Figure 2.1: DCF operation. After waiting one slot, STA2 transmits DATA to the AP, which replies with an ACK. STA1, which is a contending node, listens the ongoing transmission and defers its backoff countdown. Then, a collision is noticed because AP and STA1 end their backoff at the same time. An EIFS period is waited by the implicated nodes. Note as well that the transmitters wait for an extra slot (marked as *) before computing the new backoff.

transmit due to the sensed interference caused by other nodes, which exceeds the CCA threshold. Thus, throughput is reduced according to the number of exposed nodes that are sharing the same channel. On the other hand, the Hidden-Terminal problem may generate collisions if two out-of-range nodes transmit at the same time to a node that senses both of them. In this case, the performance is affected by a higher collisions probability, which may entail severe issues to several levels of the network architecture, such as routing or transport. Figure 2.2 shows both Hidden and Exposed Terminal Problems.

The hidden-terminal problem has been previously studied in [6, 7], which show performance loss in typical wireless deployments with moderate traffic. In order to minimise the hidden-terminal problem, the IEEE 802.11 amendment includes the optional Request-to-Send/Clear-to-Send (RTS/CTS) mechanism. It basically aims to pre-allocate the channel before performing a data transmission. To do so, both source and destination interchange RTS and CTS frames in case the channel is sensed as free from both locations. Thus, a node goes into a virtual carrier-sensing mode when listening any of the two frames, which minimises the number of potential hidden nodes. The Network Allocation Vector (NAV) is responsible to handle this virtual sensing. Figure 2.3 shows the DCF operation with RTS/CTS.

Despite the apparent enhancement provided by RTS/CTS, it may fail at solving the hidden-terminal problem. Firstly, either RTS and CTS frames may be not heard by interfering nodes. Furthermore, some counter productive effects are well-known to occur [8]. In particular, performance anomalies are likely to occur in case of existing asymmetric links, which appear from the usage of different transmit powers and carrier sense thresholds.

2.2 Scenarios

According to [9], dense scenarios will depict next-generation WLANs, so that a high number of wireless devices (e.g. 1 user/m^2) will coexist. Typical dense scenarios are provided by the IEEE 802.11ax Task Group, which are of special interest:

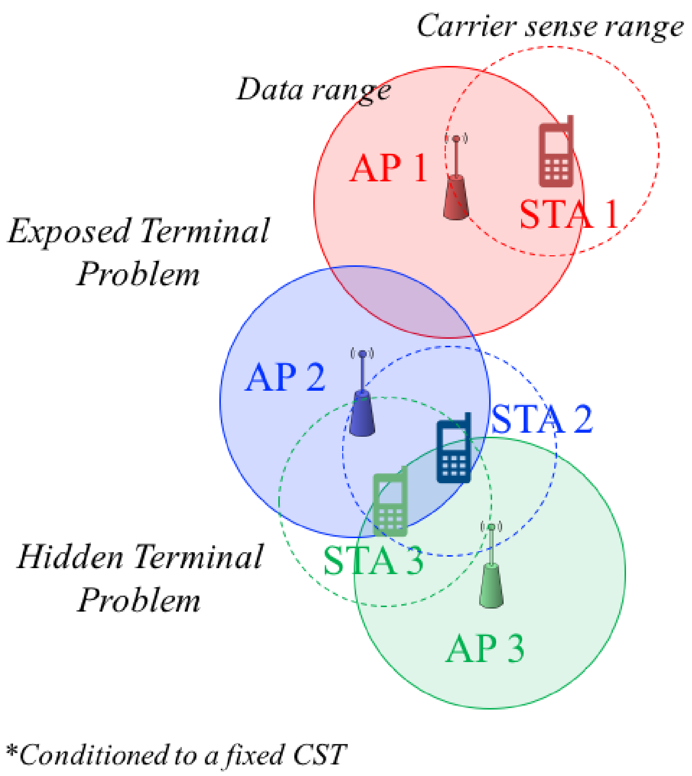


Figure 2.2: Exposed Terminal Problem: AP1 and AP2 could transmit simultaneously since STA1 and STA2 are far enough from the interference, but only a transmission at a time is allowed due to CSMA/CA. Hidden Terminal Problem: AP2 and AP3 may suffer collisions by hidden node because they are not in range of each other, but STA2 and STA3 sense both APs.

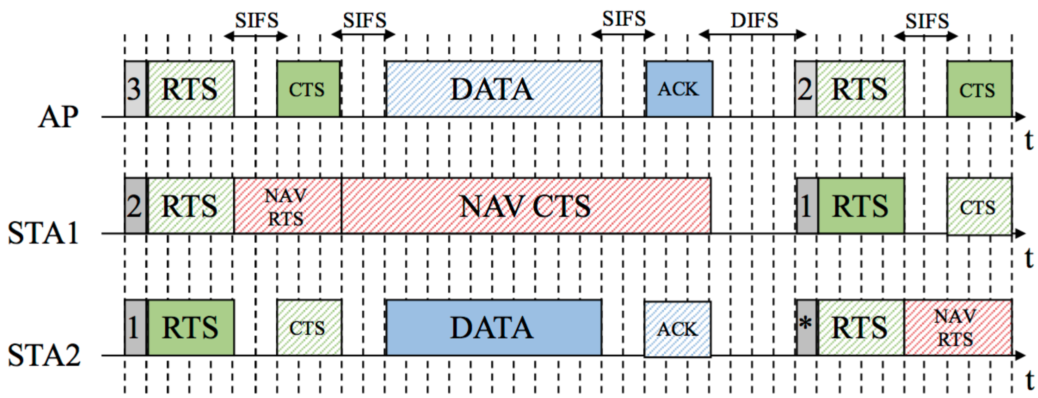
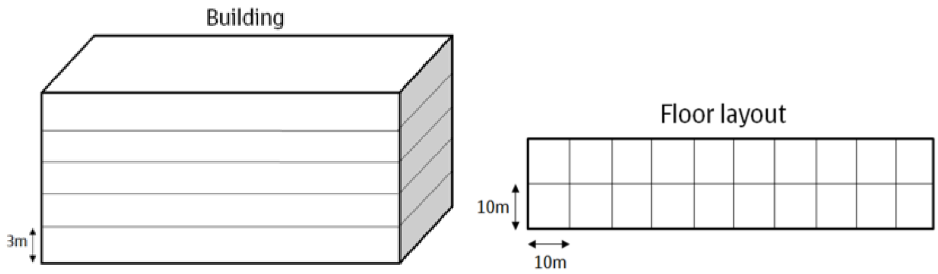


Figure 2.3: DCF operation with RTS/CTS. After waiting one slot, STA2 sends an RTS frame to indicate the channel reservation. STA1, which is a contending node, sets the NAV to perform virtual carrier sensing. In case the receiver (AP) senses the channel free, a CTS frame is replied, so that NAV is updated in STA1 to cover the entire transmission.

- **Residential Scenario:** this scenario aims to characterise a typical uncoordinated WLANs deployment, combining High-Efficiency WLANs (HEW) and legacy devices. The scenario consists in a 5-floors building of $3m$ per floor (see Figure 2.4a), containing 2×10

apartments, which have size $10m \times 10m \times 3m$ (see Figure 2.4b). An indoor path-loss model is used to capture the effect of walls and other obstacles on the propagated signals.



a Building layout b Floor layout

Figure 2.4: IEEE 802.11ax Residential scenario

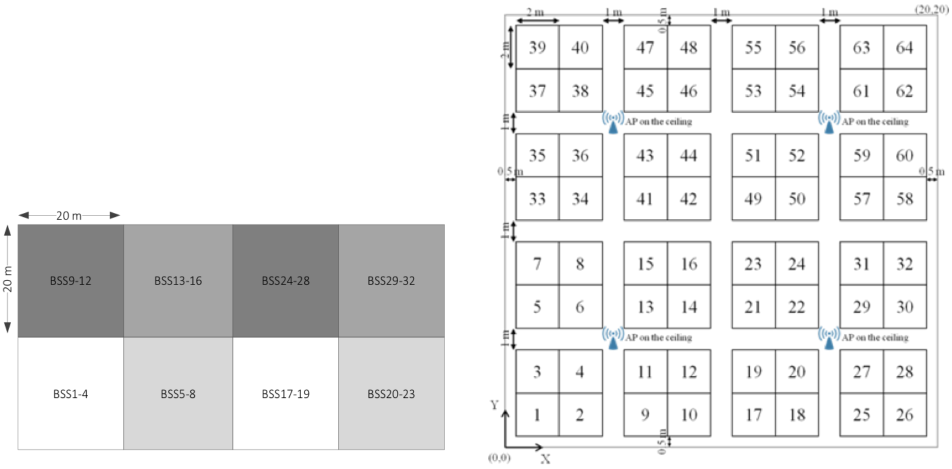
- **Enterprise Scenario:** in this case, it is presented a coordinated scenario in which most of the devices are controlled by a central entity. It is composed by 8 offices (see Figure 2.5a), containing 64 cubicles (see Figure 2.5b). In each office there are 4 APs installed in a mesh topology, and each cubicle contains 4 STAs (laptop, monitor, smartphone or tablet, and hard disk). The interference generated in such a scenario is provided by the overlapping ESSs and unmanaged networks (P2P links). The path-loss considered takes into account the walls of the offices.

- **Indoor Small BSS Scenario:** this scenario aims to represent coordinated real-world deployments that aim to extend mobile networks. To support a high number STAs, many long-range APs are placed. The ESS architecture is therefore planned through a given frequency reuse pattern (see Figure 2.6). The radius of each cell has size $10m$, and there are $2 \cdot h$ meters between consecutive BSSs, where $h = \sqrt{(R2 - R2/4)}$. The AP is placed at the centre of the cell.

- **Outdoor Large BSS Scenario:** similarly to the indoor BSS scenario, the objective for the outdoor large BSS scenario is to capture issues in real-world outdoor deployments with a high separation between BSSs. Again, the deployment is planned, so as the frequency reuse. The scenario is composed by 19 hexagonal grids, with an inter-AP separation of 130m. STAs are placed randomly through a uniform distribution, so that they associate to the nearest AP based on distance-based path loss and shadowing.

- **Residential Scenario + Outdoor Large BSS Scenario:** it is a combination of the residential and the outdoor scenarios, so that different path-loss models are considered due to walls penetration in the building case.

In addition to high density of nodes in the aforementioned scenarios, the variability in WNs in terms of user arrivals and departures entails an extra challenge for the resource allocation problem. In both [10] and [11], user arrivals are modelled as time-varying Poisson processes, such that the users arrival rate (λ) depends on the time. On these works, it is studied the users' behaviour in fully controlled environments such as a University Campus or a Conferences, but the mean time between user arrivals varies according to the scenario and many other factors.



a Offices layout b Cubicles layout

Figure 2.5: IEEE 802.11ax Enterprise scenario

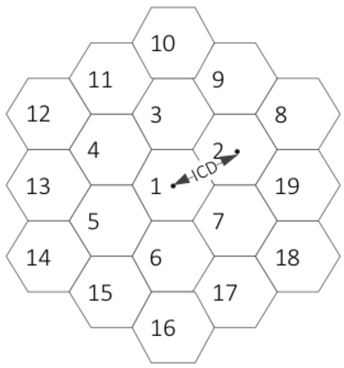


Figure 2.6: IEEE 802.11ax Cell topologies

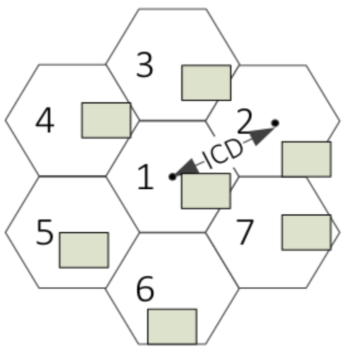


Figure 2.7: IEEE 802.11ax Residential + Large Outdoor BSS scenario

In particular, [11] considers a two-state Markov-Modulated Poisson Process (MMPP) with mean inter-arrival times of 38 seconds and 6 minutes for accessing and leaving the system, respectively.

2.3 Spatial Reuse in Overlapping IEEE 802.11 WLANs

Spatial reuse aims to efficiently use the available spectrum within a limited area, specially in dense wireless deployments. To that purpose, we find TPC, CST adjustment, and beamforming, among other techniques that are out of the scope of this Thesis (refer to Figure 1.1). For the former, by adjusting the transmit power, the number of exposed nodes can be minimised, allowing a higher number of parallel transmissions within the same range. But there is an important consideration when tuning the power transmitted, which is that data rate directly depends on the SINR sensed at the receiver (the less power transmitted, the less experienced data rate received). This has direct impact on the Modulation Coding Scheme (MCS) used. Henceforth, adjusting TPC poses the trade-off between the number of parallel transmissions and the quality of them in terms of data rate and signal robustness. Besides, TPC can help at reducing energy consumption. However, an important limitation of TPC lies in uncoordinated deployments, in which a fraction of the devices may not adjust the transmit power, thus leading to unfair configurations.

Regarding CST adjustment, the number of parallel transmissions can be also be enhanced, in addition minimising the collisions probability by hidden nodes. However, modifying the CST entails several consequences. For instance, decreasing the CST may rise probability of finding hidden nodes (the less is listened, the less is known). Table 2.1 summarises the effects of TPC and CST adaptation.

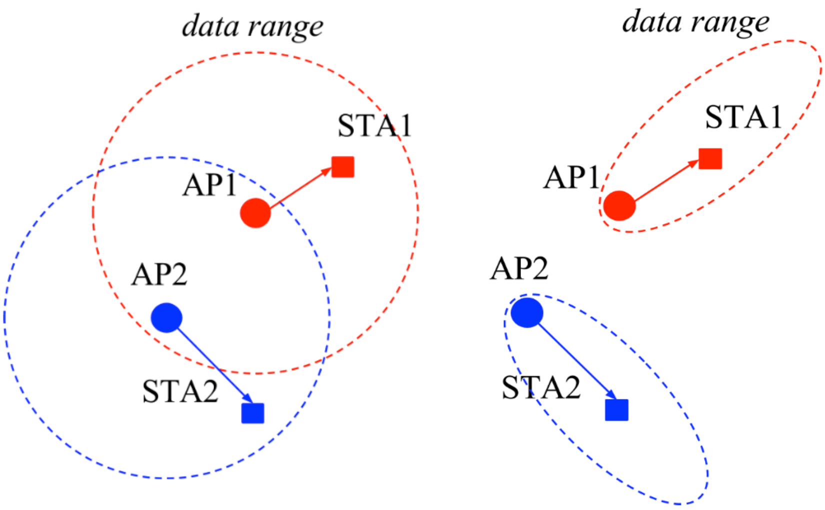
Action	**Effect**			
	Parallel Transmissions	Data Rate	Collisions probability (by hidden node)	Energy Consumption
↑ Power	↓	↑	↓	↑
↓ Power	↑	↓	↑	↓
↑ CCA	↓	-	↓	↓*
↓ CCA	↑	-	↑	↑*

Table 2.1: Effects of TPC and CST adjustment. **Adjusting CCA may indirectly affect to the energy consumption by requiring more/less power to be transmitted.*

Finally, to even improve spatial reuse we find adaptive beamforming, a technique that allows transmitting an energy beam into a specific direction. Thus, the usage of beamforming reduces the interference area for a given transmission. Figure 2.8 shows how beamforming allows enhancing the spatial reuse through directional transmissions. In 2.8a, AP1 and AP2 cannot transmit simultaneously in case of using the same frequency channel, since they are using omnidirectional antennas. However, in case of using adaptive beamforming (2.8b), the network throughput is increased due to the simultaneous transmissions from AP1 and AP2 to STA1 and STA2, respectively. The IEEE 802.11ax Task Group considers beamforming as a key technique to improve spatial reuse, which operation is defined in the V1 of the Draft [12].

2.4 Autonomous Management in Wireless Networks

For a proper operation in future WNs, we envision the usage of AI techniques to satisfy the high-demanding requirements in the complex scenarios abovementioned in Section 2.2. The



a Omnidirectional transmission b Directional transmission

Figure 2.8: Enhancement of the spatial reuse through adaptive beamforming

recent literature shows an upward trend in combining AI with Wireless Communications to solve many well-known problems such as packet routing [13], IEEE 802.11 WLAN Access Point Selection [14, 15], optimal rate sampling [16], or energy harvesting in heterogeneous networks [17]. For the particular case of modifying the transmit power and the CST, finding the optimal solution in real time is NP-hard. Additionally, adaptable and fast-convergence solutions must be granted, which entails higher complexity to find suitable solutions. For that, we put special emphasis on RL for improving the spatial reuse in high-density WNs. RL turns out to be a convenient paradigm for allowing independent devices to improve their own performance through an observations-based learning process. In RL, input signals are used to determine the behaviour of a given system, so that historical events are considered for decision-making.

Chapter 3

STATE OF THE ART

This Section describes the current literature on spatial reuse enhancement in Wireless Networks, which embraces TPC, CST adjustment and beamforming. In addition, we analyse previous work in channel allocation. Finally, since one of the main purposes of this thesis is to study the feasibility of RL application into spatial reuse enhancement, we also provide some notions on what RL is, and which kind of problems it can solve. Moreover, in this Section we introduce previous research in dynamic resource allocation in WNs through RL.

3.1 Spatial Reuse

3.1.1 Transmit Power Control

Adjusting the transmit power is one of the techniques considered in the IEEE 802.11ax standard for spatial reuse enhancement. In the first version of the draft [12], a set of regulations are provided for TPC usage. In particular, when a STA chooses a specific power detection level (OBSS_PD_{level}), the maximum transmit power is given by:

$$\text{TX\_PWR}_{max} = \begin{cases} \text{Unconstrained}, & \text{if } \text{OBSS\_PD}_{level} = \text{OBSS\_PD}_{min} \\ \text{TX\_PWR}_{ref} - (\text{OBSS\_PD}_{level} - \text{OBSS\_PD}_{min}), & \text{OBSS\_PD}_{max} \geq \text{OBSS\_PD}_{level} \geq \text{OBSS\_PD}_{min} \end{cases}$$

Despite the context provided by the TGax, modifying the transmit power is a complex task that must be carefully done, since it may affect to higher communication layers such as network (routing) and transport (congestion). Due to the fact that transmission range and the data rate are conditioned by power control at the physical layer, the main effects of increasing the transmission power in a wireless network are:

- Increases the transmission range, which is helpful for connectivity with the receiver but harmful for the overlapping nodes due to the generated interference.
- Generates more interference, which may cause a higher contention time in overlapping nodes.
- Increases the data rate.
- Increases the energy consumption.[1]

[1]Energy consumption is affected by the type of power consumption (power transmission, power idle, power sleep, power amplifier and power reception) and the devices involved in the communication. To apply a policy for saving power is strictly related to them, which is out of the scope of this Thesis.

Furthermore, TPC may create unidirectional links, which unleashes fairness issues within overlapping nodes. In fact, acknowledgements (ACKs) and RTS/CTS frames assume bidirectional links. To properly implement TPC, the first question to be asked is where should power control be done in the network architecture. This question is addressed in [18], which, in addition to devising that the network level should be in charge of power control, presents a guideline of its impact on several performance measures (end to end delay, packet losses, routing overhead, etc.). The authors claim that global optimization must be carried out by TPC approaches acting in the network layer, rather than in the MAC layer, such as many other works do. The latter only solves the TPC problem locally (usually, SINR is used to adjust power). Besides, TPC approaches acting in the MAC layer only aim to adjust power at the immediate hop, so that the next optimal hop problem is not considered (this is specially evidenced in ad-hoc WNs).

The first literature that proposes TPC is mainly focused on energy saving. Despite this topic is out of the scope of this Thesis, it is interesting to study the kind of mechanisms used and the implementations done for that purpose. The energy consumption issue is mainly approached in the context of wireless ad-hoc networks, due to the trade-off them present. On the first hand, increasing the power level increases the transmission range, but it may generate interferences that lead into collisions or starvation. On the other hand, using a low power level allows reducing the interference, but a higher number of hops may be required for a packet transmission. In [19, 20] there are presented several power control schemes relying on the RTS/CTS mechanism. In them, these management packets are used to measure the power transmitted by the overlapping wireless devices. Thus, TPC is done in such a way that the number of parallel transmissions increases whilst collisions are minimised. However, it is known that these kind of schemes can degrade the network throughput and may even incur into a higher energy consumption [21]. To solve that, [22] proposes to adjust the transmit power only during data transmissions. Thus, it is used the minimum necessary to successfully carry out the transmission, which is computed during the RTS/CTS frames exchange. Furthermore, to avoid collisions, the devices in a wireless network have to use the maximum power (p_{max}) when transmitting either an RTS or a CTS frame. In particular, RTS/CTS frames are modified to include the current and the desired power transmitted at the transmitter and at the receiver, respectively. The desired power ($p_{desired}$) is adjusted as $p_{desired} = p_{max} \times Rx_{Thresh} \times c$, where Rx_{Thresh} is the minimum necessary received signal strength and c is a constant (which in this case is set to 1). Figure 3.1 shows the behaviour of the power control scheme introduced in [22].

A further extension of this work is presented in [23], since the power control scheme shown in [22] may potentially decrease the throughput and increase energy consumption in some scenarios. This mostly occurs due to the increased collisions probability by hidden node, since either CTS or RTS frames may not be listened by some interfering nodes. Therefore, the new proposed scheme, which is named Power Control MAC (PCM), deals with the aforementioned kind of asynchronies by forcing nodes to transmit periodically DATA at the maximum power. Figure 3.2 shows the power used during a entire transmission if using PCM, which is set to p_{max} during $20\mu s$ every $190\mu s$ for the DATA case. These values are predefined because the EIFS value specified by the DCF is equal to $212\mu s$ when the data rate is 2 Mbps. Thus, power level during DATA transmission is increased once in each EIFS interval.

The PCM scheme has been validated for several network topologies, so that remarkable improvements are shown at saving energy. In particular, savings of about 10% are achieved whilst obtaining very similar throughput values with respect to the current IEEE 802.11 operation. Despite the main goal of PCM is to reduce energy consumption, it implicitly allows improving

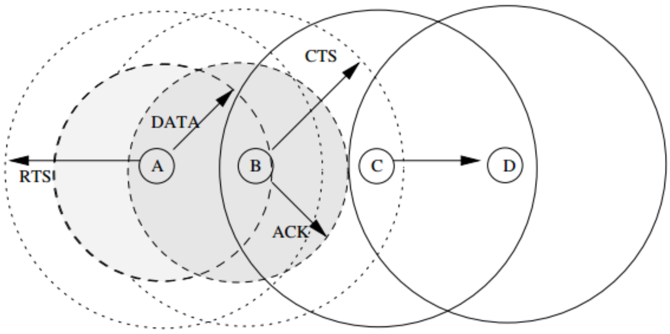


Figure 3.1: Pursley et al. Power Control Scheme. Image retrieved from [22].

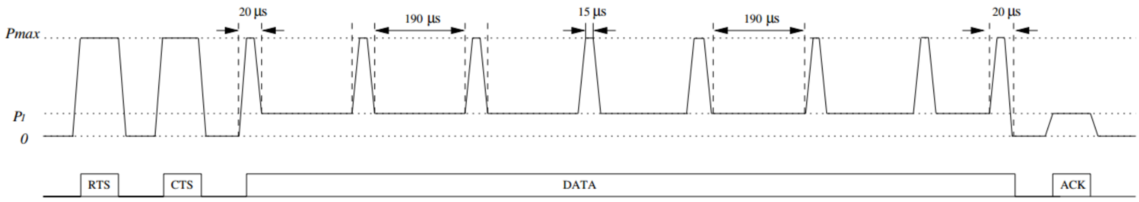


Figure 3.2: Power Control MAC (PCM) scheme. Image retrieved from [22].

the spatial reuse, which shows that using the necessary power during a wireless transmission can be beneficial for the network capacity.

Similarly to previous reviewed approaches that make use of RTS/CTS frames, the authors in [24] attempt to improve the throughput of OBSSs by using an RTS/CTS modification that incorporates different Network Allocation Vector (NAV) timers. Thus, despite RTS and CTS packets are transmitted at the maximum power, a higher number of parallel transmissions can be carried out, since the new NAV timers only captures the RTS/CTS operation. However, this solution appears to be quite rigid and to behave properly only in few situations.

Another important research line in TPC is based on measuring and/or inferring the channel characteristics. The authors in [25] show the benefits of tuning the transmit power in terms of interference, and without affecting to the MCS granted by the Channel State Information (CSI) from previous transmissions. In [26], the authors also study the potential of TPC. In particular, four TPC schemes are provided for theoretical comparison:

- No TPC: devices transmit at the maximum transmit power.
- Basic TPC: devices transmit at the minimum necessary power to reach the farthest device in their network.
- Unfiltered TPC: the power is adjusted to the minimum necessary per link communication, but beacons are set to the minimum required power for reaching the farthest device.

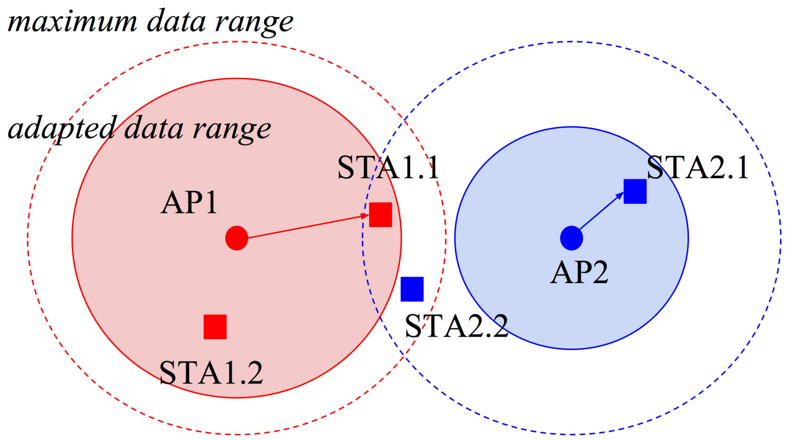


Figure 3.3: F-CMSA operation. AP1 uses a higher transmit power to reach its farthest STA (STA1.1). At the same time, AP2 transmits to a close device (STA2.1), so that transmit power does not affect the other ongoing transmission. This scheduling is provided by a central entity.

- Filtered TPC: an IIR filter is used to adjust the transmit power $y(n)$, which is computed as $y(n) = a \cdot y(n-1) + (1-a) \cdot x(n)$, where $y(n-1)$ is the previous estimated transmit power, and $x(n)$ is the instantaneous power needed.

Several experiments are done regarding the uplink and the downlink performance for several combinations of the introduced approaches, showing the potential of providing such kind of power control mechanism. In addition, it is also introduced the concept of Fractional-CSMA/CA, which aims to coordinate the transmissions within different BSSs in order to make the most of TPC. Thus, the number of parallel transmissions can be increased without suffering collisions by hidden node (refer to Figure 3.3). The concept of F-CSMA/CA is also interesting to even improve the TPC operation, since it would allow to make a step further to enhance spatial reuse. However, it requires a strong coordination that appears to be infeasible in the uncoordinated scenarios of interest for this work.

Another approach in which observations are used to determine the power level to be used is shown in [27]. In it, the authors propose a decentralised system called Dynamic Transmission Power Control (DTPC), which is based on a set of triggered thresholds. Transmission power, channel occupancy and link status are the main inputs of the algorithm for modifying the transmit power. Two simulations are carried out to validate the DTCP algorithm. The first one only considers are very simple scenario in which interferences from two APs are shown to be minimised through DTPC. The second one analyses the impact of applying DTPC in two nodes, showing improvements in both energy consumption and throughput. In both cases, adjacent channels are used. The main problem is that thresholds are set empirically (based on simulations), so it may be dangerous to apply this to other scenarios. Moreover, a centralised TPC mechanism is proposed in [28], which includes power control and Rate Adaptation (RA). An interesting contribution of this work is the creation of subgroups of WLANs according to their neighbouring relationship in terms of interference. This allows defining independent power differences between devices in the same group, which are useful to avoid asymmetric links. The main drawback of this work is that graphs can become very large, and that using the channel as a decision parameter is not strong enough to determine a relationship between two nodes.

Work	Presented Approach	Centralised/ Decentralised	Positive Aspects	Negative Aspects
[26]	Minimises transmit power while preserving the maximal rate to minimise interference	Both	Decentralised approach: • Introduces IIR filters to adjust TPC Centralised approach: • Introduces the interesting concept of F-CSMA	Decentralised approach: • Depends on rate adaptation • Only applies to dense scenarios • May generate asymmetric links • Heuristic parameters assignment is done Centralised approach: • Infeasible in uncoordinated scenarios
[24]	Presents a modification of RTS/ CTS that introduces several NAV timers, so that adjusting transmit power does not increase the number of collisions	Decentralised	• Provides fairness • Improves the network throughput	• Only useful at few specific cases. • Lack of evidences in realistic scenarios.
[27]	A real-time mechanism based on triggered thresholds is provided to modify the transmit power according to the link status and the channel occupancy	Decentralised	• No extra signalling is required	• Thresholds are set empirically • Lacks of applicability in real scenarios
[28]	Adjusts transmit power together with the data rate	Centralised	• A clustering approach is provided in order to avoid link asynchronies	• The clustering algorithm can lead to inefficient and large graphs, specially in dense environments • The cost of centralisation is not studied • Coexistence with legacy devices is not considered

Table 3.1: State-of-the-Art on Transmission Power Control

As a result, many nodes in the same group may lack of mutual interference. Similarly, the authors in [29] aim to solve the energy saving problem in ad-hoc WNs by using a protocol that dynamically adjusts the power level, so that throughput is maximised. The core idea is to create clusters, so that transmit power can be adjusted accordingly. Then, a minimum power routing scheme is provided to find the shortest path in terms of power consumption. Finally, we find a different TPC approach that is also worth to mention [30], since their authors show a strong relationship between the transmit power and the packets size at reducing energy consumption in WLANs. In particular, they adjust the transmit power according to the packet size, based on prior experiments. Furthermore, they show that the energy used to transmit one payload bit in a long run is given by:

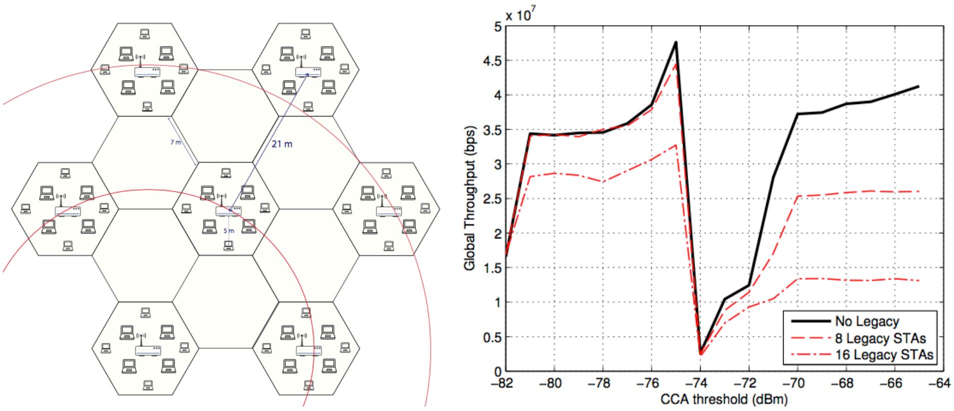
$$E_{bit_{res}} = \frac{\sum_{i=1}^{N} P_{tx,i} \cdot B_{all,i}}{B_{succ}} \cdot T_{bit},$$

where N is the number of power levels, $P_{tx,i}$ is the transmit power level i, $B_{all,i}$ is the packet size (including headers), B_{succ} is the number of successfully received bits, and T_{bit} is a bit time. As a final remark, we emphasise the importance of providing a solution that properly works in case of coexisting with legacy devices. In [31], it is argued that using TPC in such environments can be detrimental to the aggregate performance.

Table 3.1 summarises the most important reviewed works on transmit power adaptation in wireless networks.

3.1.2 Carrier Sense Threshold Adjustment

In addition to power control mechanisms, CST adjustment is key to improve the spatial reuse. By tuning the sensitivity at the receiver, it is possible to increase the number of parallel transmissions, thus maximising the aggregate throughput. In [32], the authors explore the possibil-



a Network topology b Agg. throughput vs CST value

Figure 3.4: CST modification in an IEEE 802.11 WLANs. Images retrieved from [32].

Channel width	**OBSS_PD$_{level}$**
40 MHz	OBSS_PD$_{level}$ + 3 dB
80 MHz	OBSS_PD$_{level}$ + 6 dB
160 MHz or 80+80 MHz	OBSS_PD$_{level}$ + 9 dB

Table 3.2: IEEE 802.11ax specification for the sensitivity per channel width

ities to improve the capacity in IEEE 802.11 WLANs by modifying the CST. To do so, they present a typical OBSS deployment (Figure 3.4a) for studying the network performance when using different CCA thresholds. An interesting pattern is shown regarding the CST and aggregate throughput relationship (Figure 3.4b). Moreover, the presence of legacy devices is shown to negatively affect to the overall performance.

As for TPC, the TGax also introduces several limitations to adjust the CST [12], so that minimum and maximum allowed sensitivity levels (OBSS_PD$_{min}$ and OBSS_PD$_{max}$) are determined to be -82 dBm and -62 dBm, respectively. With that, a STA can dynamically select a sensitivity level (OBSS_PD$_{level}$) during its operation under Spatial Reuse (RS) mode, which is bounded by the aforementioned minimum and maximum levels:

$$\text{OBSS\_PD}_{level} \leq max\Big(\text{OBSS\_PD}_{min}, min\big(\text{OBSS\_PD}_{max}, \text{OBSS\_PD}_{min} + (\text{TX\_PWR}_{ref} - \text{TX\_PWR})\big)\Big),$$

where TX_PWR is the STA's transmission power in dBm at the antenna connector, and TX_PWR$_{ref}$ is set to 21 or 25 dBm according to the device capabilities[2]. The obtained OBSS_PD$_{level}$ varies for each channel width used, as shown in Table 3.2.

Previous work in CST adjustment

Regarding previous work that aims to adjust the CST for spatial reuse enhancement in WNs, we first focus on the work in [33, 34], where single AP systems use the IEEE 802.11k standard

[2] TX_PWR$_{ref}$ $= 21dBm$ for non-AP STAs and for an AP with the Highest Number of Spatial Streams (NSS) Supported M1 subfield in the Tx Rx HE MCS Support field of its HE Capabilities element field equal to or less than 1. TX_PWR$_{ref}$ $= 25dBm$ for an AP with the Highest NSS Supported M1 subfield in the Tx Rx HE MCS Support field of its HE Capabilities element field equal to or greater than 2.

[35] for CST adaptation. The 802.11k defines a set of mechanisms for exchanging information between nodes in a BSS, regarding the radio channel. Thus, by implementing this standard, a node measures MAC and PHY conditions, which information is shared. In [33], an algorithm called K-APCS uses the information exchanged through the IEEE 802.11k standard to adjust the CST in real time. Motivated by the unfairness provided by K-APCS, the authors extended their work in [34], where K-APCS2 is presented. In this case, the information exchanged by nodes is used to devise problem nodes, which are targeted by the CST optimisation algorithm. A similar approach is found in [36], which uses the Packet Error Rate (PER) locally to tune the CST in a given node.

A step further for CST adaptation is found in [37], which faces the trade-off between increasing the network throughput and preserving the fairness in presence of legacy devices (i.e., devices that do not apply CST adjustment). To that purpose, the authors propose a centralised approach, Centralized Fairness Mechanism (CFM), that allow STAs to adapt their CCA according to network conditions. Accordingly, an AP implementing CFM, dynamically switches between the CCA adaptation phase, and the fixed phase. For the former, a new field is proposed to be included in beacons, which indicates if CCA adjustment is allowed or not. The decision of going into the CCA adaptation state is made by the AP according to the associated STAs, so that throughput can be maximised when no legacy devices are present.

Additionally, an interesting approach regarding CST adjustment is found in [38], which introduces the notion of non-cooperative games to represent uncoordinated environments. For that, each node maintains a utility function, defined as a non-decreasing concave function of the CST, which is used to maximise its performance. Since selfish strategies lead to an increased interference (each node wants to transmit as much as possible, so CST tends to be increased), a pricing function is defined as a constraint. A set of simulations are provided in order to show the effectiveness of their presented decentralised algorithm, in which devices update their CST every Δ seconds, according to the collisions experienced during the last interval. Results show a significant increase of the overall throughput with respect to the fixed CST strategy.

In addition to the aforementioned works in CST adjustment, we find other Multiple AP (MAP) collaborative approaches. [39, 40] propose a distributed mechanism for CST adjustment in MAP WLANs. To achieve a jointly operation, a central unit is required, which is only feasible at fully controlled environments in which the service provider has access to the entire network.

Dynamic Sensitivity Control

In addition to previous work related to CST adjustment (described in Section 3.1.2), the most important emerging approach for IEEE 802.11 WLANs is the Dynamic Sensitivity Control (DSC) algorithm, which has been proposed by the TGax [41]. The DSC algorithm aims to distributively find an appropriate level of CST in each non-AP device (i.e., an STA) of a WLAN, in order to maximise the number of parallel transmissions and, thus the aggregate throughput. DSC periodically adjusts the CST of a wireless device given the information retrieved from Received Signal Strength Indicator (RSSI) measurements. The operational mode of DSC is graphically represented in Figure 3.5.

In DSC, a station accumulates the Received Signal Strength Indication (RSSI) of each sensed beacon until the *UpdatePeriod* is reached. Meanwhile, an average of the RSSI (*AvgRSSI*) is maintained. In parallel, it is counted the number of beacons missed during a given

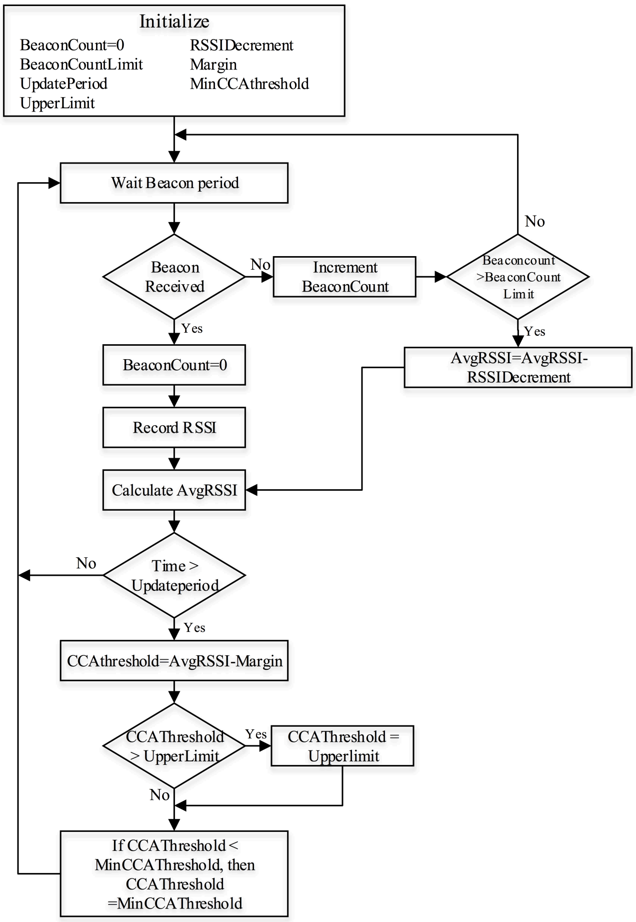


Figure 3.5: Dynamic Sensitivity Control (DSC) operational mode. Image retrieved from [42].

UpdatePeriod. At each Beacon Interval (BI), the STA compares the number of lost beacons (*BeaconCount*) with a maximum threshold (i.e., *BeaconCountLimit*). In case *BeaconCount* is greater than *BeaconCountLimit*, the average RSSI is decreased by a default value (*RSSIDecrement*) with the aim of decreasing the CST and, thus listening further away and receiving more beacons. Finally, after the *UpdatePeriod* is reached, it is computed the new CST value between *LowerLimit* and *UpperLimit*. The CST is computed as *AvgRSSI* minus a margin, which is a positive number (typically from 5 to 25).

An evaluation of the DSC algorithm is provided in [42], uses residential scenario introduced by the TGax for simulations. On the one hand, it is shown that the overall throughput can be increased up to a 20% with respect to legacy networks if using DSC. On the other hand, the application of DSC at uncoordinated WNs may generate a significant increase in the number of hidden nodes, which can potentially harm the overall throughput due to collisions. In fact, the increase in hidden nodes is measured to be the 60% in the best of the cases, and the 160% in the worst situation. Best results are obtained when using smart channel selection and fixed MCS, while the worst results are obtained when randomising both channel and modulation. DSC is also evaluated in [43], where the Indoor Small BSSs Scenario is used for simulations. In this case, setting a low margin is substantially harmful for the overall throughput due to the very high number of collisions. Moreover, for very far distances between APs, a fixed CST behaves similarly than well-performing margins for DSC. In terms of fairness, the lowest margin also provides the best results (specially in the densest case). Similar results are provided in [44, 45] for other scenarios. While the former shows DSC performance in large-scale scenarios, the latter defends that the magnitude of the gains depends on the topology, the node distribution around the AP, the distance between overlapping cells, and the presence of cell-edge nodes.

As previously seen, DSC allows enhancing the overall throughput by increasing the number of parallel transmission, and without adding any kind of signalling. However, it potentially increases the probability of suffering collisions by hidden-node, which may be counter productive if packet losses result into worse throughput performance. Furthermore, DSC may generate fairness issues, since nodes with an advantageous geographical position may have more chances for accessing the channel. Figure 3.6 shows the fairness issue in a scenario in which DSC is applied.

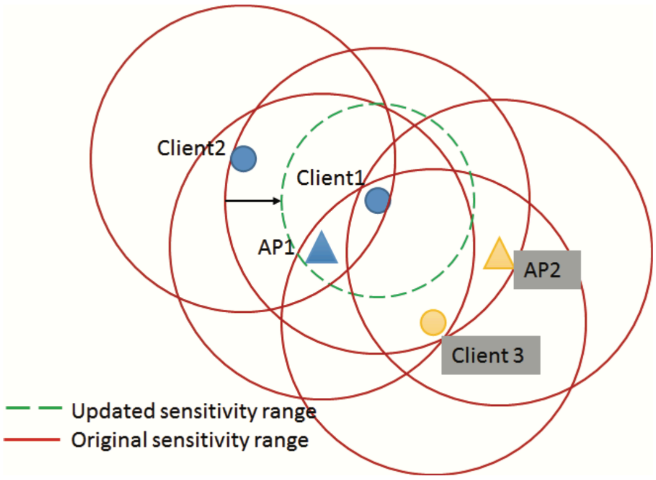


Figure 3.6: DSC Fairness Issues. Client2 should have Client1 to suspend its transmission, but DSC modifies its CST in order to dismiss Client2. Image retrieved from [43].

Further extensions of DSC have been already provided. The first one (so called DSC-AP) can be found in [46], which intends to cover both UL and DL situations by also adjusting CST on the APs. This work is motivated by the fact that the initial DSC approach does consider the starvation problem in APs during DL transmissions. Similarly to DSC, in DSC-AP the AP records the RSSI of the received frames (either data packets or ACKs) during a period called *UpdatePeriod*. The collected frames come from both associated stations and neighbouring APs. With the RSSI's information, a moving maximum RSSI (maxRSSI) is maintained for the frames received from other APs, in order to know which is the most interfering coexistent WLAN. In addition, the minimum RSSI (minRSSI) is maintained for the frames received from the STAs within the WLAN, so that the AP can detect the farthest STA (in fact, the STA from which the AP receives the poorest signal). When the *UpdatePeriod* is reached, a new CST between specific boundaries (*LowerLimit* and *UpperLimit*) is computed as:

$$\mathrm{CST}_{AP} = min(max(min\mathrm{RSSI}, max\mathrm{RSSI}) - \text{margin}, min\mathrm{RSSI}),$$

where *margin* is a value determined to be between 18 and 25 dB for indoor scenarios in which the path-loss factor (α) is 0.3. The DSC-AP algorithm is shown to improve up to a 32% the overall throughput with respect to legacy IEEE 802.11 performance. In addition, fairness is increased up to a 16% because of the reduction of exposed nodes, which supposes noticing less starvation situations. However, DSC-AP also to problems related to hidden nodes, which increase the collisions probability. But again, for the particular presented scenario, the obtained throughput is better despite the significant increase of the Frame Error Rate (FER).

To further improve the hidden nodes issues generated by DSC, another contribution of Afaqui et al. on CST adjustment ([47]) attempts to combine DSC with RTS/CTS activation in stations. In particular, three approaches for enabling RTS/CTS in a given station are shown in combination with DSC. As in previous works, the IEEE 802.11ax Residential scenario is used for simulations. The maximum improvement shows an increase of approximately 55% in the overall throughput with respect to legacy IEEE 802.11 performance. In addition, the FER is reduced almost a 40% and fairness is increased in a 30%. The fact of enabling RTS/CTS dynamically allows reducing the effects of hidden nodes in terms of collisions without adding too much overhead due to RTS/CTS transmissions, as well as it is only used in nodes that really require it. In opposite, extra communication overhead is added to decide which nodes must use RTS/CTS, which has not been quantified for showing the resulting throughput when using this approach.

To recap the content shown in this Section, a summary of the most remarkable work in CST adjustment in shown in Table 3.3.

Work	**Presented Approach**	**Centralised/ Decentralised**	**Positive Aspects**	**Negative Aspects**
[41]	Dynamic Sensitivity Control to increase the number of parallel transmissions	Decentralised	• Fully decentralised (no signalling required) • Easy for implementation • Proposed by the TGax	• It potentially increases the number of hidden nodes • Setting the *margin* parameter is not clear and does not suit different scenarios
[33, 34]	Distributed adaptive CST through IEEE 802.11k	Decentralised	• Easy to implement • Communication overhead is not an issue for the overall performance	• Lack of scalability in other scenarios • Unfairness approach
[37]	Centralised adaptive CST in presence of legacy devices	Centralised	• Provides effective results • Fair approach	• Requires a central unit • The presence of legacy nodes forces to use the default operational mode • Does not solve node problems
[39, 40]	Distributed adaptive CST for MAP WLANs	Centralised	• Provides effective results • Fair approach	• Requires a central unit • Infeasible in many scenarios

Table 3.3: Carrier Sense Threshold Adjustment State-of-the-Art

3.1.3 TPC and CCA Relationship

As previously seen in Sections 3.1.1 and 3.1.2, adjusting either the transmit power or the CST can contribute to improve spatial reuse, but presents a set of trade-offs. To even improve networks performance, there is a research line that argues for a relationship between the power transmitted and the sensitivity threshold, which is based on the naive principle "*to shout, you need to listen more carefully*". This principle is demonstrated in [48], which shows that, by using free space path loss, the optimal CST at a given receiver can be derived from the transmit power used by the transmitter. For that purpose, it first shows the relationship between the average Received Signal Strength (RSS) and the distance between the transmitter and the receiver:

$$RSS = \overline{RSS}\Big(\frac{\overline{d}}{d_{(tx,rx)}}\Big)^{\delta}$$

where RSS is the signal strength at the receiver, $\overline{RSS}$ is the power of the signal received at the defined distance $\overline{d}$ (typically 1 meter), $d_{(tx,rx)}$ is the distance between the transmitter and the receiver, and δ is the path-loss exponent (often set between 2 and 4, according to the environment). Thus, the reception range D_{tx} and the carrier sense range D_{cs} can be defined as:

$$D_{cs} = \overline{d}\Big(\frac{\overline{RSS}}{T_{cs}}\Big)^{\frac{1}{\delta}}, D_{rx} = \overline{d}\Big(\frac{\overline{RSS}}{T_{rx}}\Big)^{\frac{1}{\delta}},$$

where T_{rx} is the minimum power required at the receiver for decoding the signal and the CST, and T_{cs} is the CST. Since the correct reception of data in a given node is subject to the RSS received (must be greater than the CST) and the SINR (must be greater than the capture effect), it is worth to compute the interference range $D_{I_{tx,rx}}$ as:

$$D_{I_{tx,rx}} = d_{(tx,rx)} \times \alpha^{\frac{1}{\delta}}$$

where α is the minimum SINR required to achieve a correct decoding of the signal at the receiving node. The interference range is the greatest distance from which a node can be affected by another transmitting node. Figure 3.7 shows the relationship between D_{cs} and $D_{I_{tx,rx}}$.

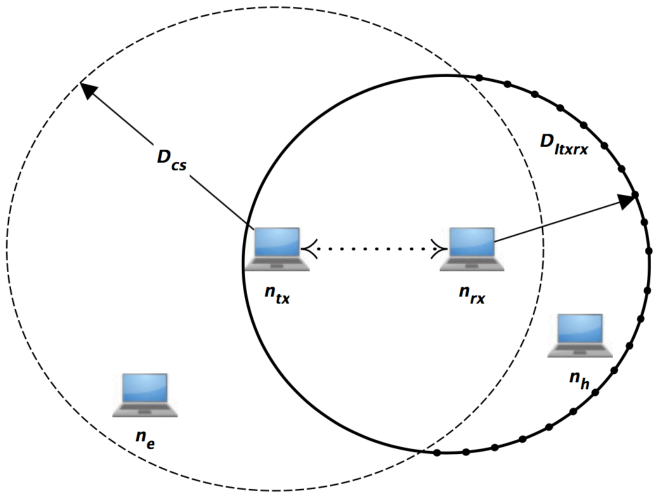


Figure 3.7: Transmit power and CST relationship. $D_{I_{tx,rx}}$ is represented by the smaller circle with the solid line, and D_{cs} is depicted by the larger circle with the broken line. Nodes n_h and n_e are hidden and exposed nodes, respectively. Image retrieved from [48].

With that, in ideal conditions, we have seen that the CST can be computed as a function of the transmit power in order to increase the spatial reuse. However, an important reason to perform CST adjustment in combination with TPC relies in the asymmetries provoked by a modification of the transmit power in an overlapping environment. The fact of creating asymmetric links generates starvation and collisions by hidden node with a higher probability, which can result into poorer performance than if using always the same configuration [18]. In relation to these performance anomalies, [49] defends that tuning TPC and CST together is a must for improving spatial reuse in WLANs. For that, it shows that links remain symmetric in a situation in which the power transmitted and the sensitivity threshold are adjusted together. Moreover, it proposes a method based on Gibbs sampling [50] for distributively adjusting both TPC and CCA. The algorithm is tested in dense scenarios (72 APs and 288 clients), showing remarkable improvements in the average throughput (it is increased by a 290%) after convergence occurs. Interestingly, from results it is observed that the power transmitted per AP is inversely proportional to the sensitivity.

Many other works exploit the relationship between the transmit power and the CST to enhance fairness and throughput performance in wireless networks. In [51], improving fairness is the main goal (although the overall throughput is also enhanced). To achieve that, each time a beacon is received at a STA, it is measured a parameter to adapt both CST and transmit power:

$$\Delta_x = Rx_p - M - \text{PCS}_{default},$$

where Rx_p is the power received (depends on the transmit power and the attenuation loss), M is a margin (in this case it is set to 20 dB), and $PCS_{default}$ is the default minimum CST (for channels of width 20 MHz, $PCS_{default}$ is -82 dBm). Then, the difference of CST and transmit power is computed from Δ_x and a given ratio:

$$\Delta_{TPC} = ratio \times \Delta_x, \Delta_{CCA} = \Delta_x - \Delta_{TPC}$$

Finally, transmit power and CST are modified by increasing/decreasing the obtained adaptation values ($TPC = TPC - \Delta_{TPC}$ and $CCA = CCA + \Delta_{CCA}$, respectively). The $ratio$ determines the adaptation degree of CST and TPC (setting the ratio to 0 entails only CST adaptation, while setting it to 1, only TPC). To both TPC and CST adaptation, so-called Balanced TPC and PCS Adaptation (BTPA), the chosen $ratio$ is 0.5, but during simulations it is also shown the results of adapting CST and TPC separately. The scenario consists in a cellular network composed by 7 APs that transmit data to 8 STAs each one. The results are significant, since the average throughput is increased by a 4 factor while preserving the fairness coefficient to 1. Despite being very simple, this mechanism allows improving the spatial reuse while preserving fairness.

The authors in [52] also consider the transmit power and CCA relationship, but put special emphasis on the trade-off between increasing spatial reuse and decreasing individual transmission data rate. Thus, they present a decentralized algorithm that adjusts power and rate control in a node, given the interference level it perceives. It is also sought to keep the highest possible data rate, while keeping minimal the generated interference towards neighbouring WLANs. To do so, the presented approach first determines the minimum power to be used so that the receiver can properly decode the message. Then, the CST is set to the minimum value at which the incoming signal is sensed, which depends on the perceived interference. To solve the problem in which the transmitter does not know the interference level I_{RX} at the receiver, the authors propose to set the CST at the transmitter as a function of the SINR. The provided adjustment function ensures that the interference at the receiver is not higher than a given threshold that depends on the power transmitted, and that the minimal data rate is sustained. The key is to restrict the transmit power based on the estimated distance between the transmitter and the interfering node. Through simulations in different scenarios (from low to high density), it is shown a significant improvement at both aggregate throughput and average saved power (up to 22% and 50%, respectively).

As in previous works, [53] shows the benefits of using both TPC and CCA adaptation in the context of IEEE 802.11ax WLANs. Furthermore, it provides evidences that different results may be obtained according to the scenario (due to path-loss effects in presence of walls). Furthermore, a rigid approach for fully controlled networks is presented in [54], where the CST to be used by devices in a mesh WN is computed as a function of the distance between nodes, and assuming that the same transmit power is used. In particular, different CST equations are provided for grid and line topologies, since the interference pattern varies. In this case, all the devices are considered to be placed symmetrically. Thus, an ideal situation is proposed for planned and interference-free deployments (e.g. a rural sensors deployment). However, this approach turns out to be unrealistic in chaotic wireless deployments.

To summarise the most remarkable work in CST adjustment in combination with TPC, in Table 3.4 we show the key features of each work reviewed in this section.

Work	Presented Approach	Centralised/ Decentralised	Positive Aspects	Negative Aspects
[49]	Gibbs sampler to both TPC and CST adjustment	Decentralised	• Dense scenarios are considered • Significant improvements in the average throughput	• Complex implementation
[51]	TPC and CCA adjustment in WLANs to improve fairness	Decentralised	• Very simple • Significant improvements in the average throughput	• Simulation scenarios are too rigid
[52]	TPC and CCA adjustment in D2D networks	Decentralised	• Very simple • Significant improvements in the average throughput	• Simulation scenarios are too rigid • The power limitation may entail performance issues to many scenarios
[54]	CCA adjustment according to the distance between nodes and the transmit power	Centralised	• Great enhancements are achieved	• Only applicable to fully controlled networks • Only rigid topologies are allowed • The same transmit power must be used by all the devices

Table 3.4: TPC and CCA Relationship State-of-the-Art

3.1.4 Beamforming

Beamforming [55] is a signal processing technique for transmitting beams of energy to a specific direction. By narrowing the signal occupancy in the space, it is possible to reduce the number of contending nodes, thus enhancing spatial reuse in an overlapping wireless scenario. Furthermore, directional antennas allow reaching farther distances, which contributes to increase the range of a network. In [56], the usage of beamforming is shown to significantly increase the network capacity, as a result of increasing spatial reuse. The application of beamforming has been mostly applied to ad-hoc wireless networks [57, 58, 59], which show great throughput enhancements since the average number of hops in packet transmissions is reduced. Nevertheless, beamforming may also lead to performance issues, such as hidden or terminal problems. According to [60], the following problems may be encountered:

- **Deafness:** occurs when a node transmits to another one that is already transmitting. As a result, the new transmission is dismissed. The deafness issue is motivated by the fact that the first transmission is not heard by the interfering node, which may also lead into a collision by hidden node. Consequences of the deafness problem in ad-hoc wireless networks are studied in [61].

- **Hidden terminal:** similarly to deafness, an ongoing transmission is not sensed, so that new transmissions may lead into collisions. The hidden terminal problem for directional antennas is studied in [62].

- **Exposed terminal:** this problem occurs when nodes are too close and the energy beam is not accurate enough to reduce the number of contending nodes. The exposed terminal problem for directional antennas is studied in [63].

- **Head of line blocking:** extra delays can occur in First-In-First-Out (FIFO) systems due to the exposed terminal problem. For instance, in case node A has packets for B and C, respectively, but B is in between an ongoing transmission, the fact of using FIFO prevents to transmit first to C, which appears to be ready for receiving the packet.

Despite of the issues arisen from the usage of directional antennas, beamforming is a promising technique that has been of particular interest in the recent years. In combination with power control, directional antennas have been shown to be effective at both throughput enhancement and energy consumption saving [64, 65].

3.2 Dynamic Channel Allocation

Despite Dynamic Channel Allocation (DCA) is contained in the frequency domain, it is worth to study it in order to relax the spatial reuse problem. DCA aims to properly allocate the available frequency range among overlapping wireless devices, so that the interference is notably reduced, and thus the aggregate throughput can be maximised. Because of the high traffic and users fluctuation experienced in wireless networks, DCA turns out to be necessary for improving the performance in a WN. Several DCA approaches can be found according to different purposes, so there is a strong discussion about the requirements that must be accomplished. While some approaches aim to provide fairness [66], some others try to grant more resources to nodes with higher traffic demands [67].

The Industrial, Scientific and Medical (ISM) band is shared by multiple technologies such as IEEE 802.11, IEEE 802.15 and Unlicensed Long Term Evolution (U-LTE). In particular, most of the Wi-Fi standards use the 2.4 GHz and 5 GHz bands, which have different characteristic. In the 2.4 GHz case, there are 14 overlapping channels of 22 MHz (refer to Figure 3.8), so that adjacent channel utilisation ($\pm$ 11 MHz) is affected by a 30 dB signal drop (shown in Figure 3.9). Similarly, the 5 GHz band is composed by three sub-bands, each one containing four non-overlapping channels. For using this band, it is required to implement mechanisms such as Dynamic Frequency Selection (DFS) or TPC.

The saturation of the unlicensed band requires a proper channel allocation, which for dense deployments becomes a challenging task due to the scarce number of available non-overlapping channels. Thus, both frequency planning and dynamic channel selection have been of remarkable interest to the wireless communications community for enhancing the throughput area in

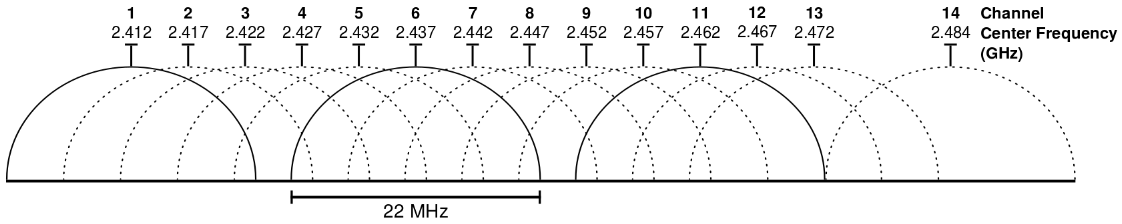


Figure 3.8: Channelisation of the 2.4 GHz band

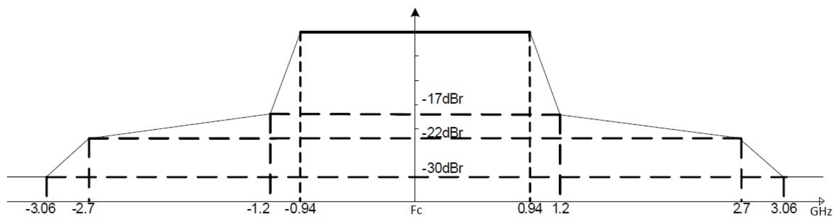


Figure 3.9: IEEE 802.11 mask

overlapping WNs. A proper frequency planning allows reducing the interference between wireless devices, thus enhancing the network capacity by avoiding collisions and delayed medium accesses. Dynamic channel selection allows wireless devices to change their centre frequency according to channel utilisation, which depends on the surrounding environment. A survey of channel assignment is provided in [68], which introduced techniques are summarised in Tables 3.5 and 3.6. Many of the reviewed techniques imply the existence of a central node that makes the channel assignment, which is infeasible for uncoordinated deployments. In contrast, decentralised mechanisms allow dynamic channel allocation is a distributed way (some of them only use local information). However, many of the decentralised approaches require from inter-AP communication and may lead to non-optimal solutions.

Centralised (or coordinated) approaches allow to effectively provide optimal (or close-to-optimal) solutions to the channel assignment problem, since complete knowledge and control on the network is provided. Moreover, in [77] it is shown that centralised approaches allow to significantly improve the performance of a network, even with the existence of independent devices. To make centralised channel allocation, the first related work is based in graph colouring techniques. In contrast, even for centralised approaches, graph colouring is an NP-hard problem, which solution computation is infeasible for large scale scenarios. To solve that, we find several works use heuristics in practical applications. The authors in [78] firstly introduced the DSATUR algorithm for that purpose, which models the frequency allocation problem using interference graphs, so that higher-degree nodes have preference at choosing their centre frequency. An extension of DSATUR is found in [79], which includes packet losses and co-channel interference in the decision-making process. In this case, the degree of a node is not given by its connections, but also by the interference and the utilisation. Several architectures are provided for running the algorithm, but all of them entail a full communication between nodes, so that centralisation is required. Another fully centralised approach in the context of wireless mesh networks is introduced by [80], showing improvements of factor 2.63 in real scenarios (wireless cards are used to run the channel assignment algorithm based on the load). In particular, the authors propose a load-aware algorithm for dynamic channel selection, so that

Technique	Description	References	With AP Placement
Graph Colouring	By assuming that APs are vertexes and non-overlapping channels are colours, graph colouring is applied to minimise the interference between adjacent cells. Graph colouring can be also used for AP placement.	[69, 70, 71, 72]	Yes / No
Integer Linear Programming (ILP)	ILP is used to place APs and assign channels by considering load balancing.	[73]	Yes
Priority-Map Approach	A floor plan is divided into pixels, which are assigned different priorities. Then, both channel assignment and AP placement is solved according to built priorities.	[67]	Yes
Patching Algorithm	An heuristic algorithm is proposed for maximising the throughput and the fairness among wireless clients, so that APs are being sequentially placed according to an objective function.	[66]	Yes
Coverage-Oriented Approach	Linear programming is used to optimise channel assignment and AP placement, so that two objective functions are used to that purposes.	[74]	Yes
Conflict-free Set Colouring	Optimal AP association is provided to minimise the interference	[75]	No
Measurement-based local coordination	Both APs and clients need to measure the interference sensed in all the frequency channels in order to derive an optimal solution for channel assignment.	[76]	No

Table 3.5: Coordinated approaches for channel assignment

traffic information is required in before-hand to provide a new channels configuration.

Centralised approaches have been shown to be very effective at maximising network performance in overlapping environments by providing a proper channel allocation. However, several limitations are found in real world. Firstly, many overlapping wireless deployments are independent, so they cannot be managed by a central entity. In addition, the required underlying communication adds a higher degree of complexity, which solution may lead to counter productive overhead. For that, we now aim to focus on decentralised approaches (with and without communication), which allow facing the channel allocation problem in a simpler way in terms of synchronisation. Additionally, for the cases in which communication is required, we are interested in quantifying the generated overhead in order to notice the actual achieved

Technique	Description	References
Least Congested Channel Search (LCCS)	APs listen to all the channels and select the one with lowest sensed interference. Other parameters such as traffic information is also used to make a decision.	[81]
MinMax Approaches	Based on the assumption that heavily loaded APs degradate the network performance, MinMax aims to minimise the maximum effective channel utilisation of those devices.	[82, 83, 84]
Weighted Colouring	From the devices' perspective, it is used the sensed interference and the number of overlapping devices to minimise an objective function that leads to channel assignment.	[85]
Pick-rand and Pick-first	Based on the sensed interference, a new channel is chosen randomly or according to a ranking list (based also on the power sensed).	[86, 87, 88]
Channel Hopping	APs change their channel periodically in order to maximise the throughput in a long run.	[89]

Table 3.6: Uncoordinated approaches for channel assignment

improvement. Regarding decentralised mechanisms for channel assignment, we first focus on [85], which presents a weighted colouring approach to improve the resource sharing problem in overlapping WNs. In particular, two techniques are provided, which are $i)$ fully decentralised and $ii)$ based on messaging passing. Again, the channel assignment problem in WLANs is approached through graph colouring, providing edges to nodes that suffer from neighbouring interference. The goal is to assign colours (i.e., channels) to nodes, so that adjacent nodes use different channels, and at the same time their usage is minimised. The fact of experiencing co-channel interference makes from this a weighted colouring problem, so an objective function is provided. This function aims to minimise the interference suffered by devices in an overlapping network. To allow decentralisation, the algorithm can only attempt to minimise the interference of a given AP, which ends up to reducing the total interference if the network progressively adjusts itself as a whole. A step further to provide fast-convergence and better results in terms of interference, is to centralise the algorithm by considering the aggregate interference. For that, a reliable communication between APs must be carried out. Notwithstanding, the implication of this communication is not even mentioned in this work, so it is not possible to determine if the presented method is practical due to the overhead or even feasible. Furthermore, [90] shows that minimising the interference through an objective function is only worth for very dense scenarios, which lacks of flexibility for maximising the throughput area. To improve it, the authors present a fully decentralized approach through a minimisation function that takes into account the total sensed interference and a cost associated to the bandwidth utilisation.

The sum of both measurements is referred as the *energy*:

$$\mathcal{E}(\mathbf{F}, \mathbf{B}) = \sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{N}_A} I_A(B) + \sum_{A \in \mathcal{A}} cost_A(b_A) \tag{3.1}$$

Where $I_A(B)$ is the interference sensed at BSS A, originated by the other WLANs $B \in \mathcal{N}_A$, and $cost_A(b_A)$ is the cost that BSS A must pay if using bandwidth b_A. The algorithm works in the APs as follows:

Algorithm 1 SAW algorithm at BSS A

Initialization:
Setup an exponential timer of mean wake-up time λ
Pick a temperature $T > 0$
Pick a random configuration $(f_A, b_A) \in \{\mathcal{F} \times \mathcal{B}\}$

When the timer fires:
Pick a random configuration $(f_{\text{new}}, b_{\text{new}}) \in \{\mathcal{F} \times \mathcal{B}\}$
Measure $\mathcal{K}_{i,A} := \sum_{B \in \mathcal{N}_A} (I_A(B) + I_B(A)) + \text{cost}_A(b_A)$, when A *does use* the configuration (f_A, b_A)
Measure $\mathcal{K}_{j,A} := \sum_{B \in \mathcal{N}_A} (I_A(B) + I_B(A)) + \text{cost}_A(b_{\text{new}})$, if A *were to use* the configuration $(f_{\text{new}}, b_{\text{new}})$
Compute

$$\beta_{ij} = \begin{cases} e^{(\mathcal{K}_{i,A} - \mathcal{K}_{j,A})/T} & \text{if } \mathcal{K}_{j,A} \geq \mathcal{K}_{i,A}, \\ 1 & \text{if } \mathcal{K}_{j,A} < \mathcal{K}_{i,A}. \end{cases}$$

Set $(f_A, b_A) = (f_{\text{new}}, b_{\text{new}})$ with probability β_{ij}
Reschedule the timer

The first experimental part (simulations carried out through Python) considers studying the modification of the cost function c/b_A (where c is a constant and b_A is the bandwidth used by BSS A), so that different dynamics of the joint performance in the overlapping networks are captured. For a null cost, the interference is attempted to be mitigated, which is shown to be effective for high density environments. The opposite occurs for a larger value of c (in this case, $c = 100$), which is only useful to improve the capacity when few overlapping networks exist, but harms the overall performance in case density increases. In addition, fairness is shown to linearly decrease as c increases. Then, a testbed implementation is provided by using 21 wireless nodes applying the channel allocation algorithm. It is shown a significant improvement for UDP and TCP traffic cases that overtakes the optimal performance achieved by applying frequency allocation through graph colouring.

Other mechanisms for decentralised channel assignment, that also rely in making measurements to enhance the network conditions, can be found in [86, 76]. [86] proposes a very simple approach in which the APs maintain an interference map of their neighbours, such that channel assignment are done through interference minimisation. Unfortunately, neither interactions between APs due to decentralisation are studied, nor strong evidences for improvements are provided. Separately, [76] proposes two decentralised approaches (in addition to a centralised one) that rely in the interference measured at both APs and STAs to calculate the best frequency channels for dynamic channel allocation. The shown uncoordinated approach presents acceptable results at minimising the interference in an overlapping network. In addition, a coordinated approach with messaging passing is presented, enhancing previous results. The main drawback of the latter is the requirement of a wired infrastructure for the management communication, which is often infeasible, specially in typical residential scenarios.

3.3 Reinforcement Learning

Reinforcement Learning frames the problem in which an agent learns by interacting with the environment (which can be dynamic), with the aim of maximising a long-term objective function. Typically, a RL problem can be modelled through a Markov Decision Process (MDP), which is defined as a tuple $\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, R, T \rangle$:

- $\mathcal{S}$ is the set of possible states in which the agent can be.
- $\mathcal{A}$ is the set of possible actions that the agent can do.
- $R_{a,s}$ is the reward generated by the environment when the agent makes an action $a \in \mathcal{A}$ when being in the state $s \in \mathcal{S}$.
- $T(s, a, s')$ is the probability of transitioning from state s to s' after doing action a.

The goal of the agent is to learn the policy (a strategy for selecting an action a in a given state s) that allows maximising the expected future cumulative reward (Figure 3.10). We find several types of rewards (episodic, average, or discounted), but for the spatial reuse problem in WNs we mainly focus on the discounted reward ($\sum_{t=0}^{\infty} \gamma^t R_{a,t}(s_t, s_{t+1})$). Note as well, that the discount factor γ allows to vary the importance of future reward, i.e., a value of γ close to 0 makes short-term reward more important, while a value close to 1 makes long-term reward more important.

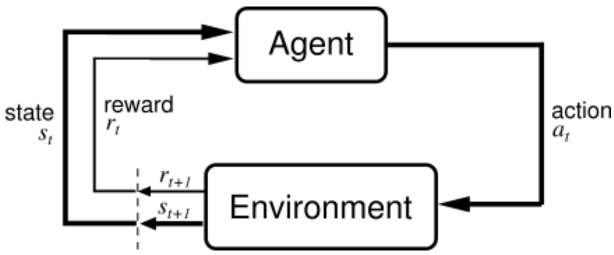


Figure 3.10: Basic RL scenario. Image retrieved from [91]

Many RL problem formulations can be found according to the desired performance (long-term or short-term reward) and the data collection process (full or partial information). The most important RL problems are modelled through Markov Decision Processes (MDPs), so that the transition between states is carried out after doing actions. Value-based methods are often used to solve MDPs, except for large-scale problems and scenarios in which full information of the model is not available. Although RL problems are often categorised into Batch (learn the entire training data) and Online learning (learn progressively), in [92] we find a nice categorisation of RL problem types and approaches, which are summarised in Figure 3.11.

The first group (Prediction), considers problems in which the value function is estimated by the learner. In practice, prediction methods describe, after ending the learning process, the expected reward when making actions from a given state. Monte-Carlo methods [93] are the basic form of value prediction, since them estimate the average value over multiple independent simulations. However, these kind of methods often perform poorly because they suffer from a high variance because of the stochastic nature of MDPs. In addition, the fact of starting from the same state results to be infeasible in many occasions. To attempt to solve the variance

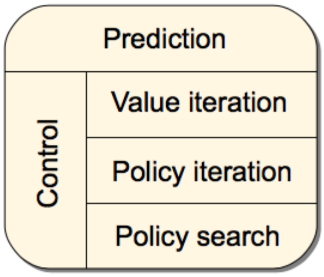


Figure 3.11: Types of RL problems and approaches. Image retrieved from [92]

problem, Temporal difference (TD) learning was firstly introduced in [94]. In opposite to the abovementioned methods, TD incorporates bootstrapping[3], so that the value estimate of the next state is used as a target for the value update. Furthermore, to maintain estimates of state values, TD includes the notion of the step-size sequence (α_t), which are non-negative numbers that regulate the importance of previous estimates in a given step t. Moreover, and to cover a large space of states (i.e., states do not fit in memory), function approximation methods are required. In this case, a feature extraction method is used to model a given state to a d-dimensional vector (e.g., with a polynomial function, state x can be converted into $\varphi(x) = (1, x, x^2, ..., x^d - 1)$). Among function approximation methods, we find gradient descend (e.g., GTD2 and TDC), and least-squares methods (e.g., LSTD(λ) and λ-LSPE).

Regarding Control methods, they aim to find a policy that maximises the reward given a set of states. To do so, we find three main approaches:

- **Value iteration:** computes the expected utility of each state by using information of neighbouring states. Thus, since value iteration is only interested in finding the optimal value, a single optimal policy is obtained if recursively applying the Bellman Equation.
- **Policy iteration:** in this approach, the following two steps are alternated:
  - A value estimate is computed for the current policy (just as for value prediction methods).
  - Then, the policy is updated such that it is greedy with respect to the new value estimate.

  Henceforth, in policy iteration, the value function is used to model the expected return of a given action-state pair, so that the immediate reward is used in combination with the discounted value of the next state (i.e., bootstrapping). Some examples of policy iteration techniques are Q-learning, TD-learning, SARSA and QV-learning. Finally, in policy search, no value function is used, but only the parametrised policy is stored. With that, it is sought to find the optimal parameters of the policy.
- **Policy search:** similarly to policy iteration, policy search methods aim to maintain the expected value for each action-state pair. However, many policy search methods rely on estimates of the expected future reward from a given state. It is mainly used in robotics, and some well-known techniques follow Episodic (e.g., Relative Entropy Policy Search,

[3]The value estimate of one state is used to update the value estimate of another state.

REPS) and Evolutionary strategies (e.g., Covariance Matrix Adaptation Evolution Strategy, CMA-ES).

Control methods can also be categorised according to the interaction that the agent maintains with the environment (see Figure 3.12). If the learner can actively influence the observations, we are in the context of interactive learning. Otherwise, with no interaction, the learner can just attempt to learn the model by only making observations.

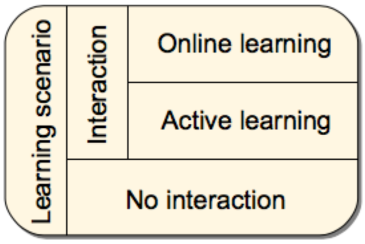


Figure 3.12: Types of RL problems and approaches. Image retrieved from [92]

In general, modelling the resource allocation problem in WNs is complex, and only partial information can be obtained. Thus, we are mostly interested in model-free learning procedures, which in fact attempt to approximate value functions. In addition, the learning procedure must be carried out in an online mode, as well as there is a high variability in WNs, specially in dense environments. For that, we envision MABs, which are further described in Section 3.3.1, to be the most suitable technique for the resource allocation problem in overlapping WNs.

Table 3.7 shows a set of RL families of methods, as well as their main characteristics.

Method	**Prediction/ Control**	**Type**	**State-dependent**	**Popular Applications**
Monte-Carlo	Prediction	Offline	Yes	Black Jack
TD	Prediction	Online	Yes	TDGammon
Q-learning	Control	Offline/ Online	Yes	Inventory management, Financial Trading Systems
SARSA	Control	Online	Yes	Gridworld, Reluctant walk
Bandits	Control	Online	No	Ranking, Resource allocation, Online advertising, Routing

Table 3.7: Families of RL methods

3.3.1 Multi-Armed Bandits

To the matter of this Thesis, we highlight the Multi-Armed Bandit (MAB) model, which turns out to properly suit to the resource allocation problem in WNs because of its characteristics. The MAB problem [96, 97] is a well-known model in the RL literature for solving resource allocation problems in which decision making must be done in face of uncertainty, but maximizing the gains. Henceforth, it attempts to characterise the exploration-exploitation trade-off.

An exploitation strategy is focused on obtaining the maximum instantaneous reward, rather than looking for a higher long-term reward, which is done by exploration. Unlike classical RL, MABs are state-independent, and it is sought to find the hidden probability distribution of the possible actions that an agent can play. The fact of not considering states allows knowing the theoretical performance limits of a given MABs problem. We find several MAB variations to model different types of problems (adversarial bandits [98], non-stochastic bandits [99], restless bandits [100], contextual bandits [101], etc.) and many different strategies for approximating a solution (ε-greedy [91], EXP3, UCB [102], etc.). Roughly, the classical Bandits problem is modelled as follows:

- Each bandit is represented by a tuple $\langle \mathcal{A}, \nabla \rangle$, where $\mathcal{A}$ is the set of arms (or actions to be chosen[4]), and ∇ is the set of probability distributions of the rewards associated to each arm, which are unknown a priori.
- After an arm $a \in \mathcal{A}$ is played at time t, a reward $r_{a,t}$ is generated by the environment.
- The goal is to minimize the regret (an indicator of the experienced losses[5]), which by definition implies maximizing the reward.
- As $t \to \infty$, the regret must be minimized as much as possible. In fact, the average regret per round becomes 0 if the action selection strategy converges, i.e., it is a *zero-regret* strategy.

Since wireless networks strongly interact with the environment, and modelling states can be extremely complex (there is a lot of information regarding each node), we mostly focus on interactive learning problems. The performance in such type of problems can be measured in many ways (sum of rewards, quantification of the mistakes done, etc.), and it is strongly related to exploration. In online learning the learner obtains data in a sequential order, so that it can be used to predict future optimal actions. Online learning is useful in case that training a entire dataset is infeasible. Moreover, it is also useful to capture dynamic environments that require self-adjusting mechanisms to perform properly over the time. State-of-the-art algorithms in online learning are mainly given for bandits (one-state MDPs) and MDPs.

3.3.2 Reinforcement Learning for improving Spatial Reuse in WNs

The current literature shows an upward trend on applying RL for solving many problems in the wireless communications field. It is on our interest to study the cases in which RL is used to improve spatial reuse in WNs, so we first focus on [103], which uses real-time Q-learning for DCA in the context of mobile communications. To do so, states for each cell are defined with the channel selected and the available channels at time t. Thus, the state of a given cell depends on the states of the others. Actions are considered to be the transition between available channels, so the reward is computed as a function that considers the channel utilisation

[4] We can indistinctly refer both to arms and to actions.

[5] The concept of *regret* is widely spread to quantify the total experienced loss with respect to the best possible action. In practice, the regret aims to counteract the environment from a game theoretical point of view, providing a notion of the learners performance. The total loss after T is given by $\sum_{t=0}^{T} \mathrm{r}_{a^*,t} - \mathrm{r}_{a,t}$, where $\mathrm{r}_{a,t}$ is the reward experienced at time t after choosing action a, and $\mathrm{r}_{a^*,t}$ is the highest reward at t provided by the optimal action a^*.

(distance between cells is considered). Despite the discount factor is deterministically set to 0.5, the learning rate is defined for each state-action pair (s, a), so that it is inversely proportional to the frequency (s, a) is visited. Q-learning is implemented distributively, so that each cell is responsible for channel assignment. However, exchanging data is required to let cells know channel-status information between neighbouring base stations. The overhead of such communication is not considered.

Q-learning is also used in [104] for channel selection in wireless networks. In this case, considerations on multi-user environments are done, so that Multi-Agent Reinforcement Learning (MARL) is applied in a scenario composed by two devices. It is shown that using Q-learning in a such competitive environment allows doing a proper distribution of the resources (in this case, frequency channels). A further extension of the previous work applying Q-learning to the channel selection problem, is presented in [105]. In this case, self-organised femtocell networks are focused to solve both carrier and power allocation problems. The former is approached through Q-learning, while the latter is formulated using the gradient method once channel allocation is done. Contrary to previous work, the authors propose a rewarding system based on the SINR. A simulation scenario with 2 cells and 5 subcarriers is used for performance analysis, which compares Q-learning with random and static strategies. Again, Q-learning parameters are deterministically set to $\alpha = 0.5$ and $\gamma = 0.9$. Another extension of this is presented in [106], which puts the focus on coordinated systems in order to improve the performance. A cooperative Q-learning approach is developed to show a higher performance and a faster convergence than uncoordinated systems.

MABs are also becoming a popular tool in the WNs field, as allow modelling the resource allocation problem without defining states. In [107], we find a distributed channel selection and transmit power approach for Device-to-Device (D2D) communications. The problem is modelled by considering $\mathcal{K} = \{1, ..., K\}$ transmitter-receiver pairs being able to choose from $\mathcal{N}'_k$ orthogonal channels and from $\mathcal{N}''_k$ power levels. The joint action selection at time t leads to the joint action profile $\mathbf{I}_t$, consisting in a pair $(\mathbf{I}'_t, \mathbf{I}''_t)$. Thus, a reward is computed for a given joint action profile $\mathbf{I}$. In order to show that the external regret of each user asymptotically decays with time, and that players actions converge into a equilibrium, two decentralized minimization strategies are presented: *No-Regret Bandit Exponential-Based Weighted Average strategy* and *No-Regret Bandit Follow the Perturbed Leader strategy*. For the former, each action is selected by some probability that depends on the accumulated reward of that action (the better performance an action shows, the higher will be its associated probability). The estimated reward is useful for calculating the regret of any other action. Basically, the algorithm makes a D2D pair to play an arm and define a so-called parameter $\delta^{(k)}_{(i \to j),t}$ according to the experienced regret (which depends on the others' actions). With that, the probability distribution of possible actions is updated, which affects to the next action. If all players play according to NR-BEWAS, then the game converges to a correlated equilibria. In the NR-BFPLS case, the action with minimum regret is chosen, but a random perturbation is added to avoid falling into a deterministic setting. Similarly to the previous approach, actions are chosen according to the current probability distribution, which is being updated at each iteration. In this case, the reward is computed by including a perturbation, which is defined as a two-sided exponential distribution. Simulations are carried out for two transmitter-receiver pairs that share two orthogonal channels. By using the two abovementioned action selection strategies, the optimal configuration is rapidly achieved by both users (approximately with 500 iterations for the NR-BFPLS strategy). Further simulations are done for a wireless network consisting in five transmitted-receiver pairs competing for three orthogonal channels with two possible power levels. The obtained reward for each of the approaches is shown to converge rapidly to the optimal.

An extension of the D2D channel selection problem through MABs is presented in [108], where Maghsudi and Stańczak include a calibrated predictor (referred in the work as *forecaster*) to infer the behaviour of the other devices. With that information, one agent is able to counter act the others' actions, thus enjoying a higher performance. In each agent, the information of the forecaster is used to choose the action with highest reward with a certain probability, while the rest of the times a random action is selected. Henceforth, assuming that all the networks use a strategy $\mathcal{X}$, fast convergence is ensured. The experimental part consists in two D2D users that must share two orthogonal channels, which shows that both users end up knowing the behaviour of the other. Thus, channel resources are optimally distributed in a very few time through a fully decentralised algorithm that does not require any kind of coordination.

Finally, a different approach for CST adaptation and TPC is presented in [109], which proposes a centralized that aims to enhancing the spatial reuse while keeping a fair distribution of the average throughput. For that purpose it is used an Artificial Neural Network (ANN) with three layers (the input layer, one hidden layer, and the output layer). The number of inputs is two times the number k of nodes to be manipulated, as well as each input represents the chosen value by each node for TPC and CCA parameters. The output in this case represents the throughput experienced by each node, so k neurons are considered.

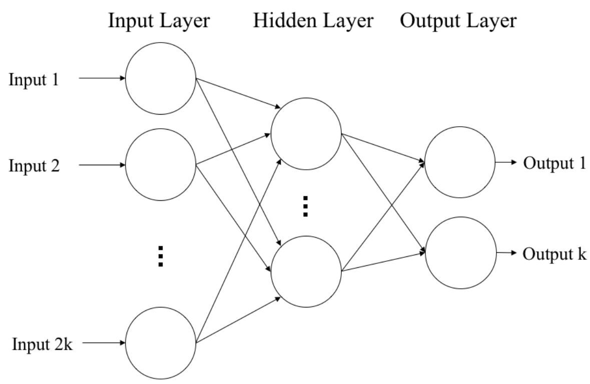


Figure 3.13: Topology of the ANN

Since the main goal is to enhance the fairness experienced in a wireless network, the minimization function used considers both performance and fairness. The authors also described the processes of data collection, training, testing and optimization, as well as the update phase. Basically, information for training/testing is constantly being updated to run the optimization, which error results are back-propagated to the previous levels. In case that the Mean Squared Error (MSE) and the desired fail number ratio (FNr$_{des}$) are below the expected results, the changes are applied and the process ends. Otherwise, both MSE and learning rate μ are decreased before running a new iteration (or epoch) of the whole process. Simulations show the performance of the ANN approach in three different scenarios: exposed-node, hidden-node, and high-density (7 APs and 8 STAs per AP). In all the cases, both aggregate throughput and fairness are shown to be enhanced after few epochs. In the high-density case, the gain noticed in the aggregate throughput exceeds the 45%, while fairness starts from 0.5 and ends up to 0.7.

Table 3.8 shows a summary of the State-of-the-art in RL applied to spatial reuse in WNs.

Work	Presented Approach	Centralised/ Decentralised	Positive Aspects	Negative Aspects
[103]	Q-learning for channel assignment in mobile communications	Decentralized	• Partially decentralised • Shows learning feasibility in WNs	• Communication overhead is not considered • Q-learning parameters are deterministically set
[104]	Q-learning for channel assignment in multi-agent environments	Decentralized	• Fully decentralised • Considers multi-agent scenarios	• Very limited scenario • Equilibrium is easy to reach
[105]	Q-learning for channel assignment in femtocell networks	Decentralized	• Fully decentralised • The rewarding system is more accurate for wireless communications	• Very limited scenario • Q-learning parameters are deterministically set
[107]	MABs for D2D channel allocation and transmit power adjustment	Decentralised	• Proves regret minimisation	• Simple scenario • Cases in which correlated equilibria cannot be reached are not studied
[108]	MABs for D2D channel allocation	Decentralised	• Interesting approach for inferring the others' behaviour	• Simple scenario • Convergence can be easily reached
[109]	ANN for adjusting the transmit power and the carrier sense threshold of all the devices in a wireless network	Centralised	• Powerful method to improve network performance	• Full control is required • Only applicable to fully centralised scenarios

Table 3.8: State-of-the-Art of Reinforcement Learning for spatial reuse in Wireless Networks

Chapter 4

CURRENT CONTRIBUTIONS AND ONGOING WORK

In this Section we describe the current contributions done during the first year, as well as the ongoing research activity and the preliminary results.

4.1 Wireless Networking Through Learning

By means of the *"Wireless Networking through Learning"* project, which is under the *María de Maeztu excellence program*, the following contributions have been done:

- A conference paper was accepted for presentation in the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) [110].[1] The work aims to study the consequences of applying RL in uncoordinated wireless scenarios. More precisely, a stateless variation of Q-learning is used to allow independent devices modifying their configuration, based on the throughput they experience. Due to the nature of the proposed scenario and the unconstrained throughput-based goal, the temporal variation of the throughput is shown to be very high. This phenomena can be understood by the adversarial setting in which actions directly affect to the others' performance.

- A journal article [111] was prepared for future submission to the Computer Networks journal.[2] Similarly to [110], the article is mainly focused in understanding the implications of applying RL in decentralised WNs. Furthermore, we generalize the contributions in [110] to different MABs implementations (ε-greedy, EXP3 and UCB). In particular, we explore the utilization of MABs when applied to the resource allocation problem in WNs, providing some guidelines for practical application. More specifically, we implement several action-selection strategies to let WNs choose the best channel and transmit power with the goal of maximizing their own throughput.

- We are actively collaborating in a journal article [112], which proposes a framework that computes Continuous Time Markov Networks (CTMNs) for throughput calculation in IEEE 802.11 WLANs [113]. With that, we are going to be able to provide realistic

[1]The article can be found at `https://arxiv.org/abs/1705.10508`.
[2]`https://www.journals.elsevier.com/computer-networks/`

results during the main research activity, since the CTMN framework allows computing the theoretical throughput for any given configuration (according to the distance between nodes, the transmit power, the CST, etc.).

- An extension of [110] and [111] is being developed for exploring centralized approaches and to provide more realistic simulations. The main purpose is to show the behaviour of a set of WNs when sharing a performance-based reward, so that different metrics can be used for experimentation (e.g., fairness, aggregate throughput, etc.). Furthermore, the previous work is extended by using the CTMN framework developed in [112], which would provide more realistic results.

- A journal article in collaboration with the Artificial Intelligence and Machine-Learning Research Group (AI-ML) is under construction [114]. The article aims to deeply study the feasibility of providing a notable enhancement of WNs performance through learning-based mechanisms. For that, several questions are addressed: i) Can learning methods guarantee optimality in this setting? ii) Are learning methods practical to wireless network optimisation? In terms of time to convergence and complexity? iii) Can the system assume performance losses due to exploration? How can we amortise those?

- A collaboration agreement was achieved with the Wireless Communications department at Universidad Autónoma Nacional de México (UNAM). The subject of this collaboration is to work in a publication that considers the inference of the coexistent devices configurations for allowing intelligent self-adjustment in dense WNs, which extends the decentralised approaches shown in [110, 111].

- A poster was presented on the 5th PhD Doctoral Workshop, held at the UPF [115].

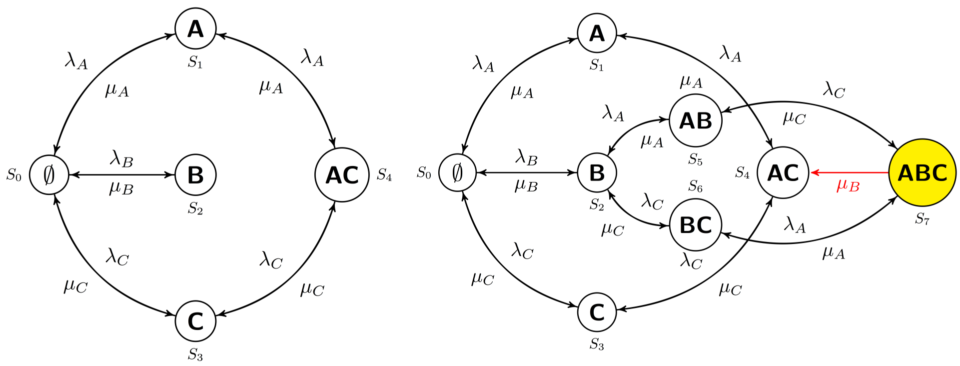
4.2 Other Projects

To support the main research activity, some other projects are being under development, which are detailed in this section.

4.2.1 Analytical Models and Simulators

A fundamental task for this Thesis is to properly understand and/or develop the most suitable simulation tools, which will allow us validating the results obtained from the research activity. One of the most important analytical models to compute the throughput in an overlapping network is the Continuous Time Markov Networks (CTMNs) model [113], which captures the operation of the CSMA/CA protocol. In a CTMN, each state is represented by the number of nodes that are transmitting, so that transitions between states are determined by the time it costs for a node to access the channel and the time it takes for leaving it, respectively. As an example, the CTMN for a scenario in which direct interference prevents a node to transmit is shown in Figure 4.1a. In particular, node B senses the channel busy if either A or C are transmitting. The throughput experienced by each node Γ_i can be directly computed as:

$$\Gamma_i = L \sum_{s \in S} \mu_{n_{i,s}} \pi_s$$



a Direct interference

b Additive interference

Figure 4.1: Markov chains for direct and additive interference schemes. Each state S_i is represented by the nodes that are transmitting simultaneously. Transitions are ruled by the duration of a packet transmission (μ_i) and the attempt rate (λ_i).

where L is the data length, $\mu_{n_{i,s}}$ is the duration of a transmission for each state s, and π_s is the fraction of time of each state, which is given by:

$$\pi_s = \frac{\prod_{i \in s} \frac{\lambda_i}{\mu_i}}{\sum_{s \in \Omega} \prod_{i \in s} \frac{\lambda_i}{\mu_i}}$$

The probability of being in state 0 (i.e., nobody transmits) is given by:

$$\pi_0 = \frac{1}{\sum_{s \in \Omega} \prod_{i \in s} \frac{\lambda_i}{\mu_i}}$$

However, this simple scenario is not likely to occur in real life, which increases the complexity of the problem for finding a solution. That is the case shown in Figure 4.1b, in which node B is prevented to transmit if both A and C transmit at the same time. A direct consequence of this on the CTMN model is that the Markov chain become irreversible, which complicates the throughput calculation. Furthermore, there appear states in which the collisions probability increases, which must be taken into account. In particular, a transitions matrix must be computed to obtain the individual throughput of each node.

In relation to CTMNs, we have started a collaboration for developing a framework that is able to build the Markov chains for analytical throughput calculation, according to the nodes input (many features such as the position, the power transmitted or the channel bonding policy are taken into consideration). The project is leaded by Sergio Barrachina [112]. Another important analytical model for throughput calculation in overlapping WLANs is the Bianchi's model [116]. The Bianchi's model is a benchmark in the current literature that we aim to use for validating the Komondor simulator (further described in Section 4.2.2).

Regarding network simulators, we highlight ns-3 [117], which is widely spread in the wireless community. However, the learning curve is high and many novel features are not yet supported. Other available simulators are ns-2 [118], OPNET [119] or Omnet++ [120].

4.2.2 Development of an IEEE 802.11ax Event-Based Simulator

In order to perform simulations based on the IEEE 802.11ax standard, we have started building our own event-based simulator for WLANs, which has been called Komondor [121]. Besides being a useful tool for supporting most of the research activity performed along the Thesis, the Komondor simulator is also the result of a collaboration with Cisco. Komondor is built on top of COST [122], which provides the baseline for synchronised messages exchanging. The decision of building a simulator, apart from the Cisco's collaboration, lies in the impediments shown by other simulators such as Network Simulator 3 (ns-3), which does not include many of the novel features provided by the IEEE 802.11ax standard. Moreover, building these functionalities in ns-3 is an extremely costly task, not only because of the complexity of the program, but for the significant learning curve. At the present time, Komondor v1.0 has been released and successfully validated against the well-known Bianchi's model for overlapping WLANs. Its main features are the Request to Send / Clear to Send (RTS/CTS) IEEE 802.11 mechanism, CSMA/CA with exponential slotted binary backoff (BEB), configurable packet traffic generation, collision type identification through logical events, and detailed statistics for metrics of interest such as average throughput per WLAN or detected hidden nodes. The code[3] is open for dissemination and improvement requests.

4.2.3 AP Association and Self-Managed Networks through Learning

We are also expecting to contribute in the collaboration Fon-UPF, which is based in two main projects for improving network performance in coexistent wireless scenarios. Fon is a company that has built the world's largest network made up of people sharing their Wi-Fi. Due to new requirements in wireless networks and the increasing density in their main deployment scenario (residential buildings), Fon aims to enhance the AP association and autonomous management processes through Machine Learning techniques. For the former, a mobile application provided to users is used to gather information about the surrounding Fon networks, whichis sent to a central unit. The latter must decide which is the best network for making user association. Secondly, to further improve networks performance, an autonomous system for management is aimed to be developed. In this case, the gathered information for decision-making is also retrieved from the APs. The core idea is to empower the system with the necessary intelligence to allow self-adjustment in Wi-Fi networks, so that the interference can be minimised, thus maximising fairness and throughput.

4.3 Preliminary Results

The research activity performed during the first year was mainly focused on understanding network interactions (emphasising dense deployments), as well as to open discussion about the potential of using RL in WNs. In [110], we applied Q-learning in a decentralised manner to the resource allocation problem in dense WNs. We showed that the best-performing actions that enhance the aggregate throughput can be found through RL. In the presented adversarial scenario (4 WNs compete for accessing to 2 overlapping channels), the most probable actions after applying Q-learning turn to be the ones that use a higher transmit power (Figure 4.2).

[3] `https://github.com/wn-upf/Komondor`

Despite resulting into a fair resources distribution (Figure 4.3a), such aggressive adversarial setting leads to a high variability in the throughput experienced by the individual networks (Figure 4.3). Hence, the favourite actions provide intermittent good/poor performance depending on the neighbouring decisions. Note that actions indexes range from 1 to 8 and, which are mapped to $\{$channel number, transmit power (dBm)$\}$ combinations: $\{1,5\}$, $\{2,5\}$, $\{1,10\}$, $\{2,10\}$, $\{1,15\}$, $\{2,15\}$,$\{1,20\}$ and $\{2,20\}$, respectively.

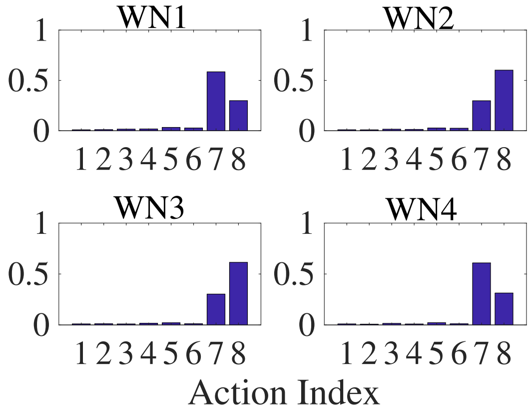
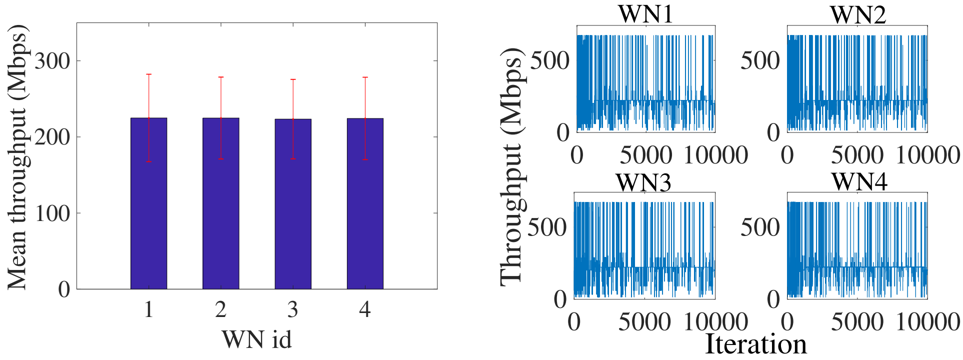


Figure 4.2: Probability of choosing an action in each overlapping WN [110]

An extension of [110] is provided in [111], in which several Multi-Armed Bandits techniques are compared to solve the same situation. Each algorithm is extensively analysed to study the actual role of each parameter (exploration, learning rate...) into network performance and interactions. Furthermore, it is shown that maximizing the individual throughput based on local information only, leads to a solution very close to the optimal proportional fairness.

Finally, to cover the centralised case, we have repeated the same experiments but using a shared reward, so that convergence can be speeded up. For that, some preliminary simulations are carried out to consider the proportional fairness[4] as the reward used by all the overlapping nodes. As shown in Figure 4.4, the variation of the temporal throughput is much lower (4.4b), at the expense of providing a lower fairness (4.4a). However, the fact of suffering a certain unfairness may result acceptable, specially in asymmetric scenarios where few nodes are in a less favourable position.

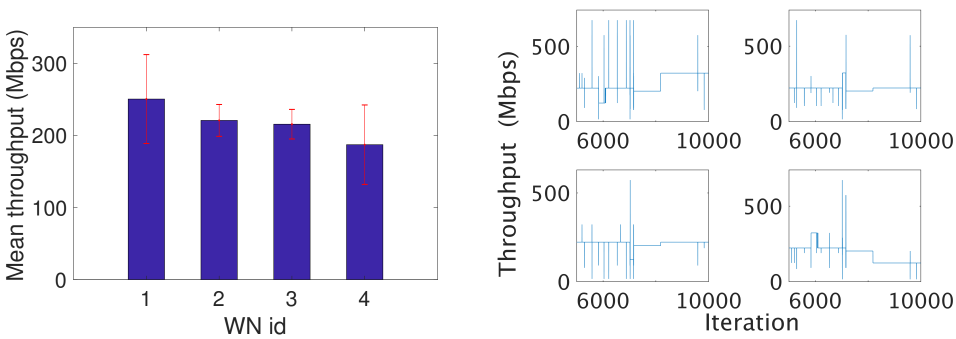
[4]The proportional fairness is the logarithmic sum of the individual throughputs of each node.



a Average throughput

b Temporal throughput

Figure 4.3: Individual performance experienced by each overlapping WN on applying decentralised Q-learning [110]



a Average throughput

b Temporal throughput

Figure 4.4: Individual performance experienced by each overlapping WN on applying centralised Q-learning

Chapter 5

PLANNING AND FUTURE WORK

We have put forward a range of contributions to be accomplished during the development of the research activities related to this Thesis, which have been introduced in Section 1.3. Here we provide further details regarding contributions, involved tasks and expected outputs.

5.1 Topic Understanding and State-of-the-Art Analysis

To properly understand the challenges and limitations of future WLANs, it is very important to analyse the current techniques in the field. In particular, we aim to focus on spatial reuse mechanisms proposed so far in WNs. Moreover, we are also expecting to deeply understand the actual behaviour and the dynamics in overlapping WNs when competing for the same resources. To do so, we are aimed to understand and/or develop analytical and simulation tools for WNs performance computation.

Involved tasks:

- Review the current literature in WNs, with special focus on spatial reuse.
- Understand current simulation tools for WNs.
- Implement new simulation tools for newest standards.
- Run experiments for coexistence analysis in WNs.

Expected output:

- A journal publication collecting these techniques to provide an State-of-the-Art to the resource allocation problem in wireless networks.
- The launching of Komondor's versions 2.1 and 2.2, which include intelligent agents to boost the network performance.
- A Komondor's technical report publication.
- A collaboration with Barrachina on the CTMN framework, which culminates in a publication.

5.2 Learning Feasibility in WNs

The next step consists in understanding the current RL techniques, in order to devise which are the ones that better suit the spatial reuse problem in WNs. Furthermore, since we are interested in competitive scenarios, we aim to study the basics on game theoretical and adversarial settings for practical application.

Involved tasks:

- Review the state-of-the-art in RL, emphasising practical application in WNs.
- Review state-of-the-art in game theory for WNs.
- Implement simulation scenarios to analyse the behaviour shown by WNs when applying RL techniques.
- Extend the work done regarding RL in WNs in order to capture dynamics in terms of user and traffic variability, which implies understanding network dynamics and future requirements for dense wireless networks.

Expected output:

- A publication of the state-of-the-art in RL applied to wireless communications.
- A publication in which RL is applied to WNs, so that convergence analysis is performed according to users variability.

5.3 Decentralised ML-based Solutions

We first aim to cover the decentralised case in wireless networks, so that ML-based solutions can be implemented in either cooperative and non-cooperative environments. For that, we aim to extend our current work in decentralised learning implications [110, 111], to explore more realistic scenarios through the CTMN framework.

In addition, we are also studying enhanced uncoordinated approaches in which WNs infer the behaviour of the coexisting devices. Thus, by using this extra knowledge, it is expected to obtain better results than for approaches using only local information.

Finally, it is expected to explore coordinated approaches for maximising the network throughput, whereby multiple devices will share a common goal (e.g., fairness maximisation).

Involved tasks:

- Apply CTMNs to current work in decentralised learning.
- Extend previous work in decentralised learning by using inference techniques that reveal the neighbours actions
- Extend previous work in decentralised learning in order to provide coordinated approaches that allow solving convergence and fairness issues.

Expected output:

- A conference paper (or a letter) regarding the application of the CTMN Framework to the work presented in [110].
- A journal publication resulting from the internship in Mexico (UNAM), which studies channel estimation and inference techniques to envision the behaviour of an overlapping WN from a decentralised point of view.
- A publication showing the trade-offs of coordinated approaches in terms of throughput enhancement and generated overhead (communication, hardware, etc.).

5.4 Centralised ML-based Solutions

Since many other typical scenarios are characterised by being composed of a central unit managing several APs, the idea of proposing centralised ML-based solutions turns out to be interesting in terms of achievable performance. Notwithstanding, it is critical to understand the trade-off between finding an optimal solution and the time it costs to be reached.

Involved tasks:

- Apply centralised ML-based techniques to the spatial reuse problem.
- Analyse the cost of centralisation in terms of communication, software and hardware overheads.

Expected output:

- A publication extending the work in decentralised learning implications [110, 111]. The work must cover the centralised point of view at devising optimal solutions.

5.5 Integration with Other Solutions

We aim to integrate the work done during this thesis with other novel solutions, so that network performance can be even improved for future deployments. More precisely, we aim to collaborate with Sergio Barrachina to merge the knowledge generated in his PhD Thesis with the one generated in ours.

Involved tasks:

- Understand work done regarding DCB.
- Integrate previous work in spatial reuse work with DCB.

Expected output:

- A publication about novel mechanisms for improving the performance in future WNs, including spatial reuse enhancement with Dynamic Channel Bonding.

Bibliography

[1] Zhenzhe Zhong, Parag Kulkarni, Fengming Cao, Zhong Fan, and Simon Armour. Issues and challenges in dense wifi networks. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2015 International*, pages 947–951. IEEE, 2015.

[2] Aditya Akella, Glenn Judd, Srinivasan Seshan, and Peter Steenkiste. Self-management in chaotic wireless deployments. *Wireless Networks*, 13(6):737–755, 2007.

[3] Jaume Barcelo, Boris Bellalta, Anna Sfairopoulou, Cristina Cano, and Miquel Oliver. Csma with enhanced collision avoidance: A performance assessment. In *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pages 1–5. IEEE, 2009.

[4] Julien Herzen, Albert Banchs, Vsevolod Shneer, and Patrick Thiran. Csma/ca in time and frequency domains. In *Network Protocols (ICNP), 2015 IEEE 23rd International Conference on*, pages 256–266. IEEE, 2015.

[5] Lu Wang, Kaishun Wu, and Mounir Hamdi. Combating hidden and exposed terminal problems in wireless networks. *IEEE Transactions on Wireless Communications*, 11(11):4204–4213, 2012.

[6] Ozgur Ekici and Abbas Yongacoglu. Ieee 802.11 a throughput performance with hidden nodes. *IEEE Communications Letters*, 12(6), 2008.

[7] Beakcheol Jang and Mihail L Sichitiu. Ieee 802.11 saturation throughput analysis in the presence of hidden terminals. *IEEE/ACM Transactions on Networking (TON)*, 20(2):557–570, 2012.

[8] João L Sobrinho, Roland De Haan, and José M Brazio. Why rts-cts is not your ideal wireless lan multiple access protocol. In *Wireless Communications and Networking Conference, 2005 IEEE*, volume 1, pages 81–87. IEEE, 2005.

[9] Boris Bellalta. Ieee 802.11 ax: High-efficiency wlans. *IEEE Wireless Communications*, 23(1):38–46, 2016.

[10] Anand Balachandran, Geoffrey M Voelker, Paramvir Bahl, and P Venkat Rangan. Characterizing user behavior and network performance in a public wireless lan. In *ACM SIGMETRICS Performance Evaluation Review*, volume 30, pages 195–205. ACM, 2002.

[11] Maria Papadopouli, Haipeng Shen, and Manolis Spanakis. Modeling client arrivals at access points in wireless campus-wide networks. In *LANMAN*, 2005.

[12] IEEE p802.11ax/d1.0, November 2016.

[13] Michael Littman and Justin Boyan. A distributed reinforcement learning scheme for network routing. In *Proceedings of the international workshop on applications of neural networks to telecommunications*, pages 45–51. Psychology Press, 1993.

[14] Biljana Bojovic, Nicola Baldo, Jaume Nin-Guerrero, and Paolo Dini. A supervised learning approach to cognitive access point selection. In *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pages 1100–1105. IEEE, 2011.

[15] Biljana Bojovic, Nicola Baldo, and Paolo Dini. A neural network based cognitive engine for ieee 802.11 wlan access point selection. In *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, pages 864–868. IEEE, 2012.

[16] Richard Combes, Alexandre Proutiere, Donggyu Yun, Jungseul Ok, and Yung Yi. Optimal rate sampling in 802.11 systems. In *INFOCOM, 2014 Proceedings IEEE*, pages 2760–2767. IEEE, 2014.

[17] Marco Miozzo, Lorenza Giupponi, Michele Rossi, and Paolo Dini. Distributed q-learning for energy harvesting heterogeneous networks. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 2006–2011. IEEE, 2015.

[18] Vikas Kawadia and PR Kumar. Principles and protocols for power control in wireless ad hoc networks. *IEEE journal on selected areas in Communications*, 23(1):76–88, 2005.

[19] Sharad Agarwal, Randy H Katz, Srikanth V Krishnamurthy, and Son K Dao. Distributed power control in ad-hoc wireless networks. In *Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on*, volume 2, pages F–F. IEEE, 2001.

[20] Phil Karn. MACA-a new channel access method for packet radio. In *ARRL/CRRL Amateur radio 9th computer networking conference*, volume 140, pages 134–140, 1990.

[21] J-P Ebert and Adam Wolisz. Combined tuning of rf power and medium access control for wlans. In *Mobile Multimedia Communications, 1999.(MoMuC'99) 1999 IEEE International Workshop on*, pages 74–82. IEEE, 1999.

[22] Michael B Pursley, Harlan B Russell, and Jeffrey S Wysocarski. Energy-efficient transmission and routing protocols for wireless multiple-hop networks and spread-spectrum radios. In *EUROCOMM 2000. Information Systems for Enhanced Public Safety and Security. IEEE/AFCEA*, pages 1–5. IEEE, 2000.

[23] Eun-Sun Jung and Nitin H Vaidya. A power control MAC protocol for ad hoc networks. In *Proceedings of the 8th annual international conference on Mobile computing and networking*, pages 36–47. ACM, 2002.

[24] Xiaoying Lei and Seung Hyong Rhee. Performance enhancement of overlapping bsss via dynamic transmit power control. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):8, 2015.

[25] Fabiano S Chaves, Andre M Cavalcante, Erika PL Almeida, Fuad M Abinader, Robson D Vieira, Sayantan Choudhury, and Klaus Doppler. Adaptive transmit power for wi-fi dense deployments. In *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pages 1–6. IEEE, 2014.

[26] Oghenekome Oteri, Pengfei Xia, Frank LaSita, and Robert Olesen. Advanced power control techniques for interference mitigation in dense 802.11 networks. In *Wireless Personal Multimedia Communications (WPMC), 2013 16th International Symposium on*, pages 1–7. IEEE, 2013.

[27] Carlos Gandarillas, Carlos Martín-Engeños, Héctor López Pombo, and Antonio G Marques. Dynamic transmit-power control for wifi access points based on wireless link occupancy. In *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pages 1093–1098. IEEE, 2014.

[28] Suhua Tang, Hiroyuki Yomo, Akio Hasegawa, Tatsuo Shibata, and Masayoshi Ohashi. Joint transmit power control and rate adaptation for wireless lans. *Wireless personal communications*, 74(2):469–486, 2014.

[29] Tamer A ElBatt, Srikanth V Krishnamurthy, Dennis Connors, and Son Dao. Power management for throughput enhancement in wireless ad-hoc networks. In *Communications, 2000. ICC 2000. 2000 IEEE International Conference on*, volume 3, pages 1506–1513. IEEE, 2000.

[30] Jean-Pierre Ebert, Björn Stremmel, Eckhardt Wiederhold, and Adam Wolisz. An energy-efficient power control approach for wlans. *Journal of Communications and Networks*, 2(3):197–206, 2000.

[31] Ioannis Broustis, Konstantina Papagiannaki, Srikanth V Krishnamurthy, Michalis Faloutsos, and Vivek P Mhatre. Measurement-driven guidelines for 802.11 wlan design. *IEEE/ACM Transactions on Networking (TON)*, 18(3):722–735, 2010.

[32] Imad Jamil, Laurent Cariou, and Jean-Francois Helard. Improving the capacity of future IEEE 802.11 high efficiency WLANs. In *Telecommunications (ICT), 2014 21st International Conference on*, pages 303–307. IEEE, 2014.

[33] Christina Thorpe, Sean Murphy, and Liam Murphy. Ieee802. 11k enabled adaptive physical carrier sense mechanism for wireless networks (k-apcs). In *Proceedings of the 4th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks*, pages 209–215. ACM, 2009.

[34] Christina Thorpe, Sean Murphy, and Liam Murphy. Ieee802. 11k enabled adaptive carrier sense management mechanism (kapcs2). In *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*, pages 509–515. IEEE, 2011.

[35] IEEE Computer Society LAN/MAN Standards Committee et al. Ieee standard for information technology telecommuni-cations and information exchange between systems-local and metropolitan area networks-specific requirements part 11: Wireless lan medium-access control (mac) and physical layer (phy) specifications. *IEEE Std 802.11ˆ¡ TM¿-2007*, 2002.

[36] Jing Zhu, Benjamin Metzler, Xingang Guo, and York Liu. Adaptive csma for scalable network capacity in high-density wlan: A hardware prototyping approach. In *Infocom*, 2006.

[37] Wessam Afifi, Enrico-Henrik Rantala, Esa Tuomaala, Sayantan Choudhury, and Marwan Krunz. Throughput-fairness tradeoff evaluation for next-generation wlans with adaptive clear channel assessment. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–6. IEEE, 2016.

[38] Kyung-Joon Park, Jennifer C Hou, Tamer Basar, and Hwangnam Kim. Noncooperative carrier sense game in wireless networks. *IEEE Transactions on Wireless Communications*, 8(10), 2009.

[39] Yao Hua, Qian Zhang, and Zhisheng Niu. A cooperative mac protocol with virtual-antenna array support in a multi-ap wlan system. *IEEE transactions on wireless communications*, 8(9), 2009.

[40] Yao Hua, Qian Zhang, and Zhisheng Niu. Distributed physical carrier sensing adaptation scheme in cooperative map wlan. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6. IEEE, 2009.

[41] G Smith. Dynamic sensitivity control-v2. *IEEE*, 802:802–11, 2015.

[42] M Shahwaiz Afaqui, Eduard Garcia-Villegas, Elena Lopez-Aguilera, Graham Smith, and Daniel Camps. Evaluation of dynamic sensitivity control algorithm for IEEE 802.11 ax. In *Wireless Communications and Networking Conference (WCNC), 2015 IEEE*, pages 1060–1065. IEEE, 2015.

[43] Zhenzhe Zhong, Fengming Cao, Parag Kulkarni, and Zhong Fan. Promise and perils of dynamic sensitivity control in ieee 802.11 ax wlans. In *Wireless Communication Systems (ISWCS), 2016 International Symposium on*, pages 439–444. IEEE, 2016.

[44] Ioannis Selinis, Marcin Filo, Seiamak Vahid, Jonathan Rodriguez, and Rahim Tafazolli. Evaluation of the dsc algorithm and the bss color scheme in dense cellular-like ieee 802.11 ax deployments. In *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2016 IEEE 27th Annual International Symposium on*, pages 1–7. IEEE, 2016.

[45] Parag Kulkarni and Fengming Cao. Taming the densification challenge in next generation wireless lans: An investigation into the use of dynamic sensitivity control. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2015 IEEE 11th International Conference on*, pages 860–867. IEEE, 2015.

[46] M Shahwaiz Afaqui, Eduard Garcia-Villegas, Elena Lopez-Aguilera, and Daniel Camps-Mur. Dynamic sensitivity control of access points for IEEE 802.11 ax. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–7. IEEE, 2016.

[47] M Shahwaiz Afaqui, Eduard Garcia-Villegas, and Elena Lopez-Aguilera. Dynamic sensitivity control algorithm leveraging adaptive RTS/CTS for IEEE 802.11 ax. In *Wireless Communications and Networking Conference (WCNC), 2016 IEEE*, pages 1–6. IEEE, 2016.

[48] Christina Thorpe and Liam Murphy. A survey of adaptive carrier sensing mechanisms for ieee 802.11 wireless networks. *IEEE Communications Surveys & Tutorials*, 16(3):1266–1293, 2014.

[49] Vivek P Mhatre, Konstantina Papagiannaki, and Francois Baccelli. Interference mitigation through power control in high density 802.11 WLANs. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 535–543. IEEE, 2007.

[50] Pierre Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media, 2013.

[51] Imad Jamil, Laurent Cariou, and Jean-François Hélard. Preserving fairness in super dense WLANs. In *Communication Workshop (ICCW), 2015 IEEE International Conference on*, pages 2276–2281. IEEE, 2015.

[52] Tae-Suk Kim, Hyuk Lim, and Jennifer C Hou. Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 366–377. ACM, 2006.

[53] Oghenekome Oteri, Frank La Sita, Rui Yang, Monisha Ghosh, and Robert Olesen. Improved spatial reuse for dense 802.11 wlans. In *Globecom Workshops (GC Wkshps), 2015 IEEE*, pages 1–6. IEEE, 2015.

[54] Jing Zhu, Xingang Guo, L Lily Yang, and W Steven Conner. Leveraging spatial reuse in 802.11 mesh networks with enhanced physical carrier sensing. In *Communications, 2004 IEEE International Conference on*, volume 7, pages 4004–4011. IEEE, 2004.

[55] Barry D Van Veen and Kevin M Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE assp magazine*, 5(2):4–24, 1988.

[56] RAMRAMAN ATHAN. Antenna beamforming and power control for ad hoc networks. *Mobile ad hoc networking*, page 139, 2004.

[57] Ram Ramanathan. On the performance of ad hoc networks with beamforming antennas. In *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*, pages 95–105. ACM, 2001.

[58] Akis Spyropoulos and Cauligi S Raghavendra. Capacity bounds for ad-hoc networks using directional antennas. In *Communications, 2003. ICC'03. IEEE International Conference on*, volume 1, pages 348–352. IEEE, 2003.

[59] Su Yi, Yong Pei, and Shivkumar Kalyanaraman. On the capacity improvement of ad hoc wireless networks using directional antennas. In *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 108–116. ACM, 2003.

[60] Basel Alawieh, Yongning Zhang, Chadi Assi, and Hussein Mouftah. Improving spatial reuse in multihop wireless networks-a survey. *IEEE Communications Surveys & Tutorials*, 11(3), 2009.

[61] Romit Roy Choudhury and Nitin H Vaidya. Performance of ad hoc routing using directional antennas. *Ad Hoc Networks*, 3(2):157–173, 2005.

[62] Aman Arora, Marwan Krunz, and Alaa Muqattash. Directional medium access protocol (dmap) with power control for wireless ad hoc networks. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 5, pages 2797–2801. IEEE, 2004.

[63] Jianfeng Wang, Yuguang Fang, and Dapeng Wu. Syn-dmac: A directional mac protocol for ad hoc networks with synchronization. In *Military Communications Conference, 2005. MILCOM 2005. IEEE*, pages 2258–2263. IEEE, 2005.

[64] Nader S Fahmy, Terence D Todd, and Vytas Kezys. Ad hoc networks with smart antennas using ieee 802.11-based protocols. In *Communications, 2002. ICC 2002. IEEE International Conference on*, volume 5, pages 3144–3148. IEEE, 2002.

[65] Asis Nasipuri, Kai Li, and Uma Reddy Sappidi. Power consumption and throughput in mobile ad hoc networks using directional antennas. In *Computer Communications and Networks, 2002. Proceedings. Eleventh International Conference on*, pages 620–626. IEEE, 2002.

[66] Xiang Ling and Kwan Lawrence Yeung. Joint access point placement and channel assignment for 802.11 wireless lans. *IEEE Transactions on Wireless Communications*, 5(10):2705–2711, 2006.

[67] P Wertz, M Sauter, FA Landstorfer, G Wolfle, and R Hoppe. Automatic optimization algorithms for the planning of wireless local area networks. In *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, volume 4, pages 3010–3014. IEEE, 2004.

[68] Surachai Chieochan, Ekram Hossain, and Jeffrey Diamond. Channel assignment schemes for infrastructure-based 802.11 wlans: A survey. *IEEE Communications Surveys & Tutorials*, 12(1), 2010.

[69] Alex Hills. Large-scale wireless lan design. *IEEE Communications Magazine*, 39(11):98–107, 2001.

[70] Petri Mahonen, Janne Riihijarvi, and Marina Petrova. Automatic channel allocation for small wireless local area networks using graph colouring algorithm approach. In *Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. 15th IEEE International Symposium on*, volume 1, pages 536–539. IEEE, 2004.

[71] Janne Riihijarvi, Marina Petrova, and Petri Mahonen. Frequency allocation for wlans using graph colouring techniques. In *Wireless On-demand Network Systems and Services, 2005. WONS 2005. Second Annual Conference on*, pages 216–222. IEEE, 2005.

[72] Janne Riihijarvi, Marina Petrova, Petri Mahonen, and Jovander De Almeida Barbosa. Performance evaluation of automatic channel assignment mechanism for ieee 802.11 based on graph colouring. In *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, pages 1–5. IEEE, 2006.

[73] Youngseok Lee, Kyoungae Kim, and Yanghee Choi. Optimization of ap placement and channel assignment in wireless lans. In *Local Computer Networks, 2002. Proceedings. LCN 2002. 27th Annual IEEE Conference on*, pages 831–836. IEEE, 2002.

[74] Andreas Eisenblatter, Hans-Florian Geerdes, and Iana Siomina. Integrated access point placement and channel assignment for wireless lans in an indoor office environment. In *World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a*, pages 1–10. IEEE, 2007.

[75] Arunesh Mishra, Vladimir Brik, Suman Banerjee, Aravind Srinivasan, and William A Arbaugh. A client-driven approach for channel management in wireless lans. In *Infocom*, 2006.

[76] Jeremy K Chen, Gustavo De Veciana, and Theodore S Rappaport. Improved measurement-based frequency allocation algorithms for wireless networks. In *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*, pages 4790–4795. IEEE, 2007.

[77] Akash Baid and Dipankar Raychaudhuri. Understanding channel selection dynamics in dense wi-fi networks. *IEEE Communications Magazine*, 53(1):110–117, 2015.

[78] Daniel Brélaz. New methods to color the vertices of a graph. *Communications of the ACM*, 22(4):251–256, 1979.

[79] Eduard Garcia Villegas, Rafael Vidal Ferré, and Josep Paradells. Frequency assignments in ieee 802.11 wlans with efficient spectrum sharing. *Wireless Communications and Mobile Computing*, 9(8):1125–1140, 2009.

[80] Ashish Raniwala, Kartik Gopalan, and Tzi-cker Chiueh. Centralized channel assignment and routing algorithms for multi-channel wireless mesh networksrancesc. *ACM SIGMOBILE Mobile Computing and Communications Review*, 8(2):50–65, 2004.

[81] Murali Achanta. Method and apparatus for least congested channel scan for wireless access points, October 5 2004. US Patent App. 10/959,446.

[82] Kin K Leung and B-J Kim. Frequency assignment for ieee 802.11 wireless networks. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 3, pages 1422–1426. IEEE, 2003.

[83] Ming Yu, Hui Luo, and Kin K Leung. A dynamic radio resource management technique for multiple aps in wlans. *IEEE Transactions on Wireless Communications*, 5(7):1910–1919, 2006.

[84] Ming Yu and Hui Luo. An adaptive radio resource management technique for aps in wlans. In *Networks, 2004.(ICON 2004). Proceedings. 12th IEEE International Conference on*, volume 1, pages 85–91. IEEE, 2004.

[85] Arunesh Mishra, Suman Banerjee, and William Arbaugh. Weighted coloring based channel assignment for wlans. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(3):19–31, 2005.

[86] Robert Akl and Anurag Arepally. Dynamic channel assignment in ieee 802.11 networks. In *Portable Information Devices, 2007. PORTABLE07. IEEE International Conference on*, pages 1–5. IEEE, 2007.

[87] Mohamad Haidar, Robert Akl, Hussain Al-Rizzo, and Yupo Chan. Channel assignment and load distribution in a power-managed wlan. In *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, pages 1–5. IEEE, 2007.

[88] Hussain Al-Rizzo, Mohamad Haidar, Robert Akl, and Yupo Chan. Enhanced channel assignment and load distribution in ieee 802.11 wlans. In *Signal Processing and Communications, 2007. ICSPC 2007. IEEE International Conference on*, pages 768–771. IEEE, 2007.

[89] Arunesh Mishra, Vivek Shrivastava, Dheeraj Agrawal, Suman Banerjee, and Samrat Ganguly. Distributed channel management in uncoordinated wireless environments. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 170–181. ACM, 2006.

[90] Julien Herzen, Ruben Merz, and Patrick Thiran. Distributed spectrum assignment for home WLANs. In *INFOCOM, 2013 Proceedings IEEE*, pages 1573–1581. IEEE, 2013.

[91] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[92] Csaba Szepesvári. Algorithms for reinforcement learning. *Synthesis lectures on artificial intelligence and machine learning*, 4(1):1–103, 2010.

[93] Christian P Robert. *Monte carlo methods*. Wiley Online Library, 2004.

[94] Richard S Sutton. Learning to predict by the methods of temporal differences. *Machine learning*, 3(1):9–44, 1988.

[95] Yotam Abramson and Yoav Freund. Active learning for visual object recognition. *Technical report, UCSD*, 2004.

[96] Peter Auer, Nicolo Cesa, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

[97] Michael N Katehakis and Arthur F Veinott Jr. The multi-armed bandit problem: decomposition and computation. *Mathematics of Operations Research*, 12(2):262–268, 1987.

[98] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Foundations of Computer Science, 1995. Proceedings., 36th Annual Symposium on*, pages 322–331. IEEE, 1995.

[99] Satyen Kale, Lev Reyzin, and Robert E Schapire. Non-stochastic bandit slate problems. In *Advances in Neural Information Processing Systems*, pages 1054–1062, 2010.

[100] Peter Whittle. Restless bandits: Activity allocation in a changing world. *Journal of applied probability*, 25(A):287–298, 1988.

[101] Naoki Abe, Alan W Biermann, and Philip M Long. Reinforcement learning with immediate rewards and linear hypotheses. *Algorithmica*, 37(4):263–293, 2003.

[102] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

[103] Junhong Nie and Simon Haykin. A q-learning-based dynamic channel assignment technique for mobile communication systems. *IEEE Transactions on Vehicular Technology*, 48(5):1676–1687, 1999.

[104] Husheng Li. Multi-agent q-learning of channel selection in multi-user cognitive radio systems: A two by two case. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, pages 1893–1898. IEEE, 2009.

[105] Mehdi Bennis and Dusit Niyato. A q-learning based approach to interference avoidance in self-organized femtocell networks. In *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, pages 706–710. IEEE, 2010.

[106] Mehdi Bennis, Sudarshan Guruacharya, and Dusit Niyato. Distributed learning strategies for interference mitigation in femtocell networks. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–5. IEEE, 2011.

[107] Setareh Maghsudi and Sławomir Stańczak. Joint channel selection and power control in infrastructureless wireless networks: A multiplayer multiarmed bandit framework. *IEEE Transactions on Vehicular Technology*, 64(10):4565–4578, 2015.

[108] Setareh Maghsudi and Sławomir Stańczak. Channel selection for network-assisted D2D communication via no-regret bandit learning with calibrated forecasting. *IEEE Transactions on Wireless Communications*, 14(3):1309–1322, 2015.

[109] Imad Jamil, Laurent Cariou, and Jean-François Hélard. Novel learning-based spatial reuse optimization in dense WLAN deployments. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):184, 2016.

[110] Francesc Wilhelmi, Boris Bellalta, Cristina Cano, and Anders Jonsson. Implications of decentralized Q-learning resource allocation in wireless networks. *arXiv preprint arXiv:1705.10508*, 2017.

[111] Francesc Wilhelmi, Boris Bellalta, Jonsson Anders Neu Gergely Cano, Cristina, and Sergio Barrachina. Enhancing spatial reuse in wireless networks through decentralised multi-armed bandits: Possibilities, challenges and perils. *arXiv preprint arXiv:1705.10508*, 2017.

[112] Bellalta Boris Barrachina, Sergio and Francesc Wilhelmi. Dynamic spectrum access in high-density wlans. 2017.

[113] Boris Bellalta, Alessandro Zocca, Cristina Cano, Alessandro Checco, Jaume Barcelo, and Alexey Vinel. Throughput analysis in csma/ca networks using continuous time markov networks: a tutorial. In *Wireless Networking for Moving Objects*, pages 115–133. Springer, 2014.

[114] Cano Cristina Jonsson Anders Gmez Vicen Neu Gergely Barrachina Sergio Wilhelmi Francesc Bellalta, Boris and Maddalena Nurchis. Learning from experience in future wireless networks: Big challenges and opportunities. 2017.

[115] Francesc Wilhelmi, Boris Bellalta, Cristina Cano, and Anders Jonsson. Improving spatial reuse in high-density wireless networks through learning. *5th EITIC Doctoral Workshop*, 2017.

[116] Giuseppe Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3):535–547, 2000.

[117] Gustavo Carneiro. Ns-3: Network simulator 3. In *UTM Lab Meeting April*, volume 20, 2010.

[118] Teerawat Issariyakul and Ekram Hossain. *Introduction to network simulator NS2*. Springer Science & Business Media, 2011.

[119] OPNET Modeler. Opnet technologies inc, 2009.

[120] András Varga and Rudolf Hornig. An overview of the omnet++ simulation environment. In *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*, page 60. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

[121] S. Barrachina-Muñoz and F. Wilhelmi. Komondor network simulator. *GitHub repository*, 2017. [Online] Available: `https://github.com/wn-upf/Komondor` (commit).

[122] Gilbert Chen, Joel Branch, Michael Pflug, Lijuan Zhu, and Boleslaw Szymanski. Sense: a wireless sensor network simulator. *Advances in pervasive computing and networking*, pages 249–267, 2005.