

Throughput Optimization Strategies for Large-Scale Wireless LANs

Mostafa Pakparvar*, David Plets*, Jeroen Hoebeke*, Dirk Deschrijver*, Michael Mehari*,
Tom Dhaene*, Ingrid Moerman*, Luc Martens* and Wout Joseph*

*Ghent University/iMinds, Gaston Crommenlaan 8, Bus 201,9050 Gent, Belgium
Email: David.Plets@intec.ugent.be

Abstract—Thanks to the active development of IEEE 802.11, the performance of wireless local area networks (WLANs) is improving by every new edition of the standard facilitating large enterprises to rely on Wi-Fi for more demanding applications. The limited number of channels in the unlicensed industrial scientific medical frequency band however is one of the key bottlenecks of Wi-Fi when scalability and robustness are points of concern. In this paper we propose two strategies for the optimization of throughput in wireless LANs: a heuristic derived from a theoretical model and a surrogate model based decision engine.

I. INTRODUCTION: BACKGROUND AND OBJECTIVES

Robustness is a fundamental requirement of many wireless networks which serve as the infrastructure for crucial services. There are many parameters that put the robustness of wireless networks at stake. If we define the robustness of wireless networks in terms of continuous connectivity and preserving a certain level of QoS, all parameters that break these two conditions should be investigated in order to maintain the robustness of the network.

Spectrum shortage jeopardizes performance of wireless networks. Most of the current wireless standards operate on the unlicensed industrial scientific medical band where the open sharing model [1] is the only dynamic spectrum access mechanism. All wireless terminals compete for the same spectrum with no hierarchical (authoritative) ascendancy. Many techniques are used by the medium access control (MAC) layers to facilitate co-existence of the links on the same frequency band. For instance, the CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) mechanism of the IEEE 802.11 standard is designed to tackle the homogeneous interference -which is caused by the same type of network-efficiently. However, this type of spectrum sharing mechanism imposes an upper bound on the number of users sharing the same channel. The availability of several IEEE 802.11 non-overlapping channels at both the 2.4 and 5 GHz bands, increases the probability of inefficient spectrum utilization for networks which are not aware of the wireless environment and operate on a fixed channel. Spectrum white spaces have already been utilized in standards like IEEE 802.22 in the TV bands where vacant frequency bands are reused by secondary users who continuously probe the existence of primary (licensed) users in order to avoid interfering them.

Also mobility of wireless devices influences both connectivity and QoS degradation conditions. If the network is not

properly planned, poor coverage introduces disconnections and poor QoS at certain locations. Moreover, mobility introduces uncertainty in the number of clients associated to any of the access points since clients are free to move and hence making hand-overs to different access points. As such, it is likely that too many clients end up in the same collision domain where the QoS requirement level is not guaranteed to be met for that number of users.

There are many intelligent channel switching algorithms in literature that address the aforementioned issues for both wireless sensor networks and wireless LANs [2], [3]. In [4] the authors propose a framework, which utilizes radio environment maps to make informed decisions on the operating channel of the wireless network in response to performance degradation of the network due to the interference. However, the framework was verified in a controlled pseudo-shielded environment with a limited number of devices.

This paper presents two possible ways for throughput optimization in a wireless environment with multiple interfering users: a theoretical model based decision engine and a surrogate model based decision engine.

II. DESCRIPTION OF THE ENVIRONMENT

The wireless network is located in an open indoor environment (e.g., industrial manufacturing hall, conference hall) covered by several Access Points (APs) dividing the area into a number of cells (see Figure 1). Like many large-scale industrial wireless networks, the mobility of the clients is addressed by making hand-overs to corresponding APs upon moving between the cells. A set of external ad-hoc Wi-Fi links are considered as the sole interference sources to the main network.

During the network operation, several fixed monitoring nodes will capture and analyse the present Wi-Fi traffic to characterize both the main network traffic as well as the interference activity.

III. THEORETICAL MODEL

A. Architecture of the framework

As shown in Figure 2, the decision making process is triggered when network failure occurs i.e., when any of the clients fails to meet its target performance metric (throughput) requirement for more than a specific number of times. We assume that the network failure event is triggered either when

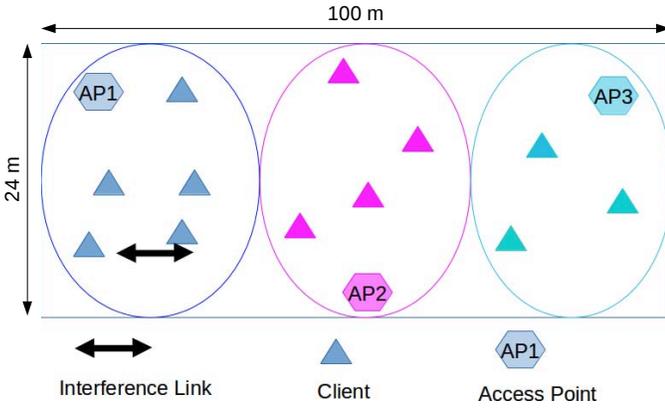


Fig. 1. The configuration of the wireless environment which consists of three cells (solid line oval regions) inside an industrial manufacturing hall.

the Wi-Fi channel has to be shared among too many competing users or when low transmission rate of one of the users imposes the performance anomaly phenomenon [5] to the whole cell. The former may occur due to the free movement of the clients within different cells. Nonetheless, the Wi-Fi performance anomaly is unavoidable when e.g. an external co-channel interference transmits at lower bit rates or when a client, in response to high packet error rates, transmits at lower bit rates governed by the automatic rate control mechanism of the IEEE 802.11 standard [6].

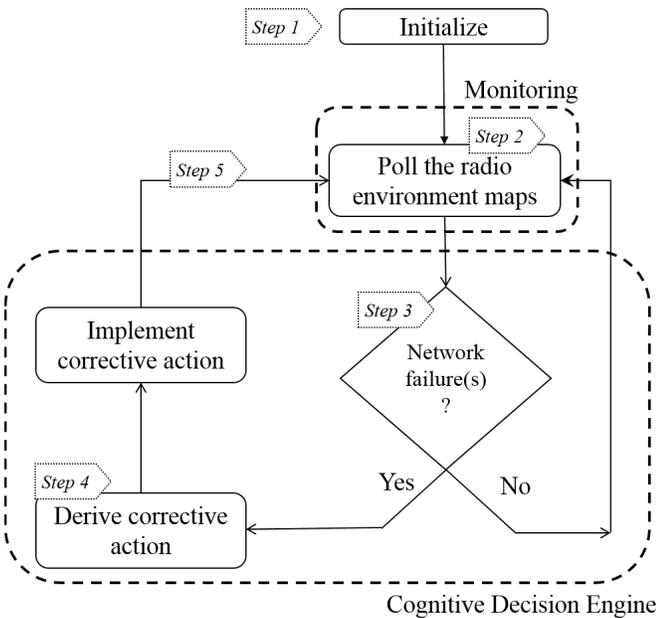


Fig. 2. Architecture of the framework.

To show the efficiency of this framework, we define a sequence of actions. When the framework is initialized (step 1), all clients start periodic iperf throughput tests to the experiment controller. The radio environment maps [7] at this stage show that network is performing normally (step 2). A network failure event is triggered on cell x . The maps show

the performance degradation of clients in cell x (step 3). The cognitive decision engine (CDE) switches the channel of AP x to the lowest risky channel if any is found (step 4). The performance improvement is shown on the maps (step 5). The CDE considers all the possible actions and their risks to survive the network.

B. Heuristic Decision Making Algorithm

The CDE uses a heuristic algorithm to find the best configuration of the network. The inputs of the algorithm are the following parameters.

- Number of clients associated to each AP
- Data generation rate of each client
- Transmission rate of each client

As output the algorithm switches the channel of the APs. If the CDE decides to switch the channel of one of the APs due to a network failure event, it should take into account if this decision introduces capacity issues to the neighboring cells if the target channel is overlapping with any of the neighboring cells. The CDE should also consider the cost of switching the channel of an AP since in this case the clients associated to the target AP have to switch their channels or make hand-overs, which might not be interesting for some use cases.

The theoretical ground based on which, we develop our decision making algorithm is the analytical model of the IEEE 802.11 DCF (Distributed Coordination Function) performance described firstly in [5] and extended further in [8]. The model estimates the useful achievable throughput of several contending stations by deriving the overall transmission time of the packets as well as the approximation of collision probability. The overall transmission time for every packet is the summation of payload and overhead transmission times plus the time spent on contending other stations for accessing the spectrum. The model is derived with the assumption that there are $N - 1$ stations transmitting at R Mbps and the single other station transmits at a lower rate of r Mbps. All stations are connected to the same Access Point (AP) while their transmission buffer is always full, i.e. they always have packets to transmit towards a server behind the AP. The model suggests that after a sufficient time, the throughput of all the stations would be almost equal although it is intuitively expected to obtain higher throughput on terminals with bit-rate of R .

Since the original model was developed for the IEEE 802.11b standard, extending the same model to newer technologies requires correct technology dependent values for SIFS (Short Interframe Space), DIFS (DCF Interframe Space), and the slot time when calculating the overhead times of the PLCP (Physical Layer Convergence Protocol) header and preamble. Once the appropriate values are looked up for calculating the overhead times, the eventual throughput X on the i^{th} channel is known by

$$X_i(r, R, N) = \frac{s_d}{(N - 1)T_f + T_s + P_c(N) \times t_{jam} \times N} \quad (1)$$

where s_d is the data frame length in bits, T_f and T_s represent packet transmission time for the fast and slow terminals

respectively, $P_c(N)$ is the proportion of collisions experienced for each packet successfully acknowledged at the MAC level ($0 \leq P_c(N) < 1$), and finally t_{jam} is the average time spent on collisions. The reader is referred to [5] for further details. Figure 3 illustrates the realization of model in Equation 1 for $R = r$ while Figure 4 illustrates the realization for $R \neq r$ and $r = 11$ Mbps. Both figures suggest that the performance of the IEEE 802.11g standard is at stake when the number of contending stations exceeds a few. Moreover, equal probability of channel access for all stations regardless of their bit-rate, ends in unfair distribution of channel access time and hence the so-called performance anomaly [5] of IEEE 802.11.

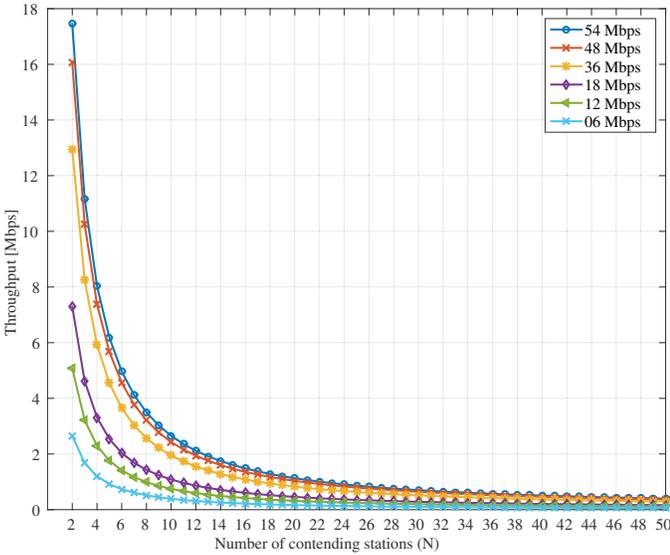


Fig. 3. Realization of the throughput model for $R=r$.

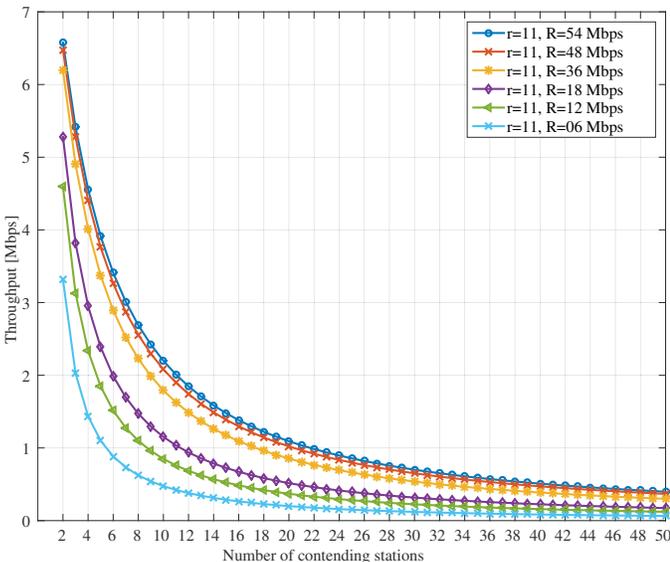


Fig. 4. Realization of the throughput model for various R values and $r=11$ Mbps.

Using the model in Equation 1, the CDE is able to estimate the performance of the network for every possible decision, i.e. switching the channel of any of the APs to any of the available channels. In this way, some of the decisions might introduce high collision probability since a large number of terminals might end up in the same collision domain (channel). Therefore, by ranking the outcome of every decision in terms of estimated throughput and attributing a cost to each decision, the CDE easily finds the best channel of operation for the right AP by:

$$i^{\text{best}} = \arg \min_i \left[\omega_i^{N_i} \mid X_i(r, R, N_i) \geq X_{\min} \right] \quad (2)$$

where $\omega_i^{N_i}$ denotes the corresponding cost of switching to channel i with N_i contending stations, and X_{\min} is the user-defined guaranteed throughput of the network. The generic approach we have used here enables the user to define custom relations for the decision cost ω_i .

IV. SURROGATE MODELLING

A. Architecture

The architecture of the surrogate model based decision engine mainly consists of the predictive surrogate model and the optimization algorithm, as illustrated in Figure 5.

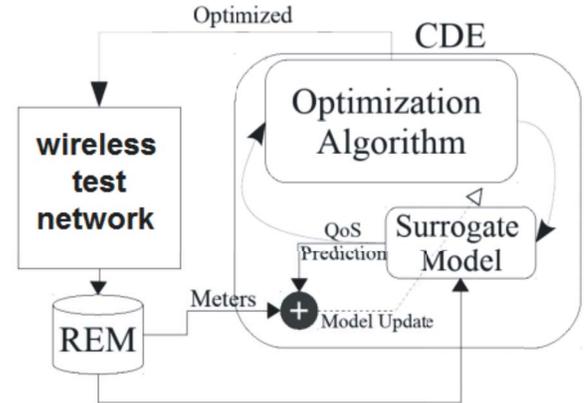


Fig. 5. Architecture of the surrogate modelling system.

The System Under Test (SUT) is the wireless system that needs to be optimized. The Radio Environment Map (REM) actively monitors the environment through spectrum sensing, network sniffing, ... and combines all monitoring information into a map of measurable parameters (MP). The Cognitive decision engine (CDE) is a self-learning decision engine that is based on QoS optimization by taking a suitable cognitive action in response to the REM information. Based on the result of each action, the algorithm assesses the new performance of the system (QoS). The surrogate model is a black-box model that characterizes and mimics the behavior of the wireless system. It takes the current state of the network (MP parameters) and the current configuration of the wireless system (Controllable parameters (CP)) as input and makes a prediction

of the QoS. The algorithm tunes CP parameters of the wireless system with the help of the surrogate model predictions. The outcome of the optimization is used to improve the accuracy of the model predictions during its operation. In the initial phase of the algorithm, the model learns the behavior of the systems by making exploratory decisions and observing the outcome. During the deployment phase it deploys knowledge to make better decisions.

B. QoS Model

The QoS model is formulated as a compact mathematical expression that can be used to forecast instantaneously how the QoS will be influenced for all different values of the knobs (CP parameters) for different values of the meters (MP parameters). In order to obtain some initial knowledge of the wireless system, the algorithm goes through a learning phase where a limited set of $k = 1, \dots, K$ experiments are performed. Each experiment k is defined by a vector of $i = 1, \dots, n_c$ knobs $C_k = C_k^i$ that are configured by the decision engine and a vector of $i = 1, \dots, n_m$ meters $M_k = M_k^i$ which are continuously monitored from the REM. The outcome of each experiment is represented by a tuple: $C_k, M_k, \text{QoS}(C_k, M_k)$. The experiments that are performed lead to a set of data samples that are used to build a surrogate model, predicting the QoS metric (throughput) for the experiments that have not yet been performed (i.e. different values of C and M), based on a distance-based similarity measure. As additional experiments are performed, the history database grows over time and the model will be able to make more accurate predictions.

C. Optimization Algorithm

Once a surrogate model is available, the optimization algorithm monitors how the meters M change over time, and automatically tunes the knobs C to maximize the predicted QoS. The values of the meters M from the REM are an input for the surrogate model, and a set of candidate settings for the knobs C is generated, to evaluate the surrogate model and see which setting yields an optimal QoS metric value. The surrogate model is built with a suitable set of data samples such that the algorithm has sufficient knowledge to make reliable predictions of the QoS. The collection of the data samples happens in 2 phases: in the beginning, the algorithm focuses mainly on *exploration* because it is learning the behavior of the system. As time progresses, knowledge is built up and the algorithm focuses more on *exploitation* where it deploys its knowledge.

The initial exploration phase tries different settings of the knobs C in order to cover the design space as evenly as possible. Once this exploratory setting of the knobs is applied on the system, the resulting QoS value will be evaluated and the outcome is added as an additional data sample to the history database. As the environment changes, the meters will vary dynamically over time and additional data samples will be added to the history database leading to more accurate predictions of the surrogate model.

As the algorithm learns and acquires more data samples, the next step is to exploit the learned knowledge in order to propose optimal decisions instead of performing merely exploratory decisions. To this end, one can use the accurate surrogate model along with a simple optimization to generate these settings. Just like in the exploration phase, the meters are monitored and a large set of candidate settings is generated for the knobs in such a way that they cover the entire control space.

For all the candidate settings of the knobs C , the surrogate model is evaluated and a setting C_{opt} is chosen for which the model predicts the highest QoS value. Hence, the algorithm searches for a candidate setting of the knobs C for which the predictive model gives the best QoS values.

Once the so-called optimal setting C_{opt} of the knobs is applied on the system, the resulting QoS value is evaluated and the outcome is compared to the predicted QoS value of the surrogate model. If the difference between these two is larger than a predefined threshold, then the prediction of the model was not sufficiently accurate and the data sample is added to the history database. If the surrogate model prediction was sufficiently accurate, then the action of the decision engine is considered successful and no further actions are needed. As the environment changes, the meters M will vary over time and additional data samples will be added to the history database, hereby leading to more accurate predictions of the surrogate model.

D. Practical experiment

A practical case study is conducted in a pseudo-shielded testbed environment w-iLab.t [14] with WiFi/Zigbee nodes. The setup that was considered is a typical scenario of two IEEE 802.11 standard compliant nodes generating traffic (SUT, node 50 to 51, see Figure 6). At the same time, two external interference links (INT, 28 to 40, 39 to 49) are activated.

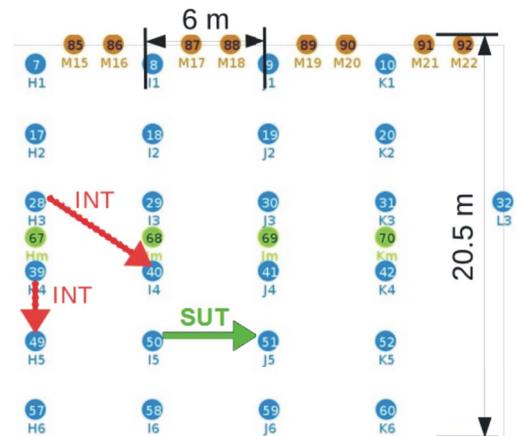


Fig. 6. Testbed with indication of the system under test and the interference links.

In this scenario, the proposed CDE optimizes the throughput of the SUT. We assume that INT and SUT operate on the same or non-overlapping IEEE 802.11g channels [16]. In this condition, the throughput of the SUT is influenced by the channel occupancy degree (COD) of the interference link. Thus, the meters in this scenario are the COD values of IEEE 802.11g channels 1 and 6, the knob is the SUT channel of operation (1 or 6). The INT links operate on channel 1 and 6 with interference CODs according to each experiment configuration. Over time, the COD of the INT links is varied after a certain number of time steps during learning. Each time step is configured at 30 seconds, which allows for reliable throughput values in presence of a fixed interference. The CDE will react and switch between optimal operating points. The COD of CH 1 and 6 are meters, the SUT channel is a knob. The decision engine monitors the current value of the meters and optimizes the channel of SUT, in such a way that the SUT throughput is maximized. As the meter changes over time, the CDE adjusts the optimal knob (SUT channel) instantaneously. Since a real-time response of the CDE is desired, the engine is subjected to a learning phase where it learns the behavior of the system.

During the learning phase, an initial set of 87 experiments are performed to learn the resulting throughput over the SUT link as a function of 3 parameters (interference COD on channels 1 and 6, SUT operating channel). The meters change in an uncontrollable way. However, the values of the knob are chosen by the CDE in such a way that the parameter space is well covered, i.e. a balanced trade-off is made between exploitation (trying configurations that are most likely to give good throughput performance) and exploration (trying configurations that are different from what that the algorithm has tried before) of the design space. Using these selected samples, an interpolation model is generated by the CDE so as to find a surrogate model for the throughput over the entire design space. The model obtained from the selected set of learning measurements is successfully validated.

Finally, the performance of the CDE is assessed by comparing the obtained throughput in 15 experiments with different configurations. Once with an static operation of the SUT and second time with utilization of the CDE to control the SUT channel. The results are shown in Figure 7 where for each time step, its corresponding meter and knob values are also indicated. From time step 1 to 6, channel 6 was the optimum channel of operation for the SUT since as seen in Figure 7 staying on channel 1 caused a huge performance degradation for the case where the CDE was not deployed. The CDE switched the SUT channel at time steps 6 when according to its internal surrogate model, a better throughput was predicted. Figure 5 also reveals the performance improvement of the CDE at different time steps when different interference configurations are in place, e.g. at the worst case condition (time step 1) more than 100% performance improvement was achieved.

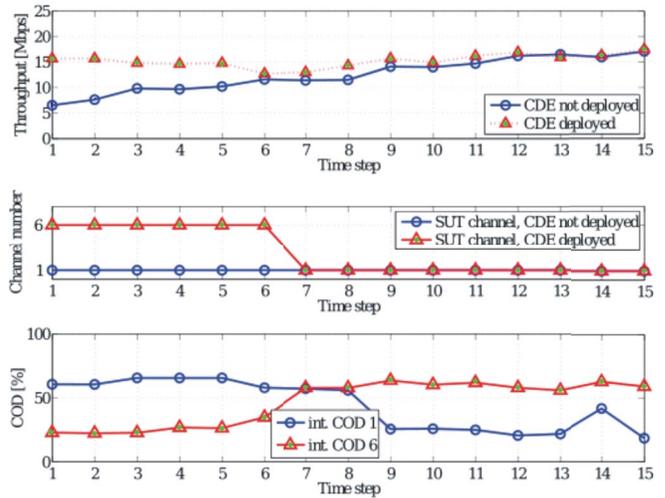


Fig. 7. Throughput performance with(out) CDE, SUT CH number, and INT COD.

V. CONCLUSION

Spectrum scarcity is imposing serious issues for the unlicensed band wireless networks necessitating adoption of cognitive radio techniques in such networks. We have proposed two frameworks that incorporate radio environment maps for throughput optimization: one heuristic based on a theoretical model and one decision engine based on surrogate modelling. A practical application example is described for the latter, showcasing its usability. In this way the performance of the clients in a large scale wireless network can be optimized. Future work consists of a thorough comparison of the two methods.

ACKNOWLEDGMENT

This work was supported by the iMinds-ICON QoCON project, co-funded by iMinds, a research institute founded by the Flemish Government in 2004, and the involved companies and institutions. D. Deschrijver is a Post-Doctoral Fellow of the FWO-V (Research Foundation - Flanders). The research is also partly funded by the Fund for Scientific Research - Flanders (FWO-V, Belgium) project G.0325.11N.

REFERENCES

- [1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.
- [2] M. Pakparvar, H. Gharibdoust, S. Pollin, and L. Tytgat, "Dynamic channel selection algorithms for coexistence of wireless sensor networks and wireless lans," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2013 IEEE 9th International Conference on*, pp. 33–38, IEEE, 2013.
- [3] L. Tytgat, B. Jooris, P. De Mil, E. De Poorter, I. Moerman, and P. Demeester, "Multichannel protocol for interference avoidance in wireless sensor networks," in *2nd Gent University and KEIO University Global COE Joint workshop 2008*, pp. 55–58, 2008.
- [4] M. Pakparvar, D. Plets, E. Tanghe, D. Deschrijver, W. Liu, K. Chemmangat, I. Moerman, T. Dhaene, L. Martens, and W. Joseph, "A cognitive QoS management framework for WLANs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–18, 2014.

- [5] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2, pp. 836–843 vol.2, March 2003.
- [6] "IEEE standard for information technologyTelecommunications and information exchange between systems local and metropolitan area networksSpecific requirements part 11," *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pp. 1 –2793, 2012.
- [7] Y. Zhao, *Enabling Cognitive Radios through Radio Environment Maps*. PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, May 2007.
- [8] A. Duda *et al.*, "Understanding the performance of 802.11 networks.," in *PIMRC*, vol. 8, pp. 2008–1, 2008.