A New Self-planning Methodology Based on Signal Quality and User Traffic in Wi-Fi Networks

Juan Sánchez-González, Jordi Pérez-Romero, Oriol Sallent

Universitat Politècnica de Catalunya (UPC), c/ Jordi Girona, 1-3, 08034 Barcelona, Spain
juansanchez@tsc.upc.edu

Abstract. Wi-Fi networks have become one of the most popular technologies for the provisioning of multimedia services. Due to the exponential increase in the number of Access Points (AP) in these networks, the automation of the planning, configuration, optimization and management tasks has become of prime importance. The efficiency of these automated processes can be improved with the inclusion of data analytics mechanisms able to process the large amount of data that can be collected from Wi-Fi networks by powerful monitoring systems. This paper presents a new self-planning methodology that collects historical network measurements and extracts knowledge about user signal quality and traffic demands to determine adequate AP relocations. The performance of the proposed AP relocation methodology based on a genetic algorithm is validated in a real Wi-Fi network. The proposed approach can be easily adapted to other contexts such as small cell networks.

Keywords: Self-planning, Wi-Fi, genetic algorithms, data analytics.

1 Introduction

In the last years, mobile users are increasingly demanding new multimedia services (i.e. high quality video, online multimedia applications, augmented/virtual reality, etc.) that have high bandwidth demands and strict Quality of Service (QoS) requirements. To cope with this demand, network densification through the deployment of small cells operating cellular technologies (e.g. 4G/5G), complemented with Wi-Fi hotspots, which benefit from the unlicensed use of the spectrum, becomes a relevant solution. Indeed, the popularity of Wi-Fi technology among mobile users makes it a competitive option for serving multimedia demands. The amount of traffic of IEEE 802.11 (i.e. Wi-Fi) has suffered a high increase in the last years. It is expected that by 2021, 63% of the global cellular data traffic will be offloaded to Wi-Fi or small cell networks [1]. Globally, there will be nearly 549 million public Wi-Fi hotspots by 2022, up from 124 million in 2017, a fourfold increase. Such a high increase in the amount of traffic volume and the number of Wi-Fi Access Points (AP)/Small Cells makes much more complex the planning, configuration, optimization and management tasks of these networks. For this reason, the automation of these processes is of paramount importance. Legacy systems such as 2G/3G/4G already started the path towards a higher degree of automation through the introduction of Self-Organizing

Network (SON) functionalities [2], namely self-planning, self-optimization and self-healing. However, the lessons learnt from cellular SON cannot be directly applied to Wi-Fi systems. Several works such as [3][4] identify challenges and use cases to introduce SON paradigms in Wi-Fi networks. Concerning self-planning aspects in Wi-Fi/Small Cells networks, most of the works found in the literature focus on the determination of the most adequate AP location and the frequency channel allocation and AP transmitted power. As an example, [5] presents a framework for automated cell-planning in multi-tenant Small Cells networks that considers actions such as adding/removing channels and Small Cells. Similarly, [6] presents a theoretical model to determine the most adequate AP location by an accurate characterization of impact of the environment (walls, doors, etc.) on the signal propagation. Other works, such as [7], deal with the determination of the most appropriate locations of Wi-Fi extenders.

On the other side, it is envisaged that the efficiency of automated network management processes can be substantially enhanced through the exploitation of powerful data analytics technologies able to process the large amount of data that can be collected from Wi-Fi networks by powerful monitoring systems. In this respect, (big) data monitoring and analytics combined with SON technologies will become fundamental in order to enable full scale automated network deployment and optimization [8]. The monitoring system provides the ability to collect information about the network resources and performance of the services while the analytics system allows extracted knowledge of the collected data in order to support different decision making processes over the network. As an example, [9] presents a framework for data monitoring and analytics for the optimization of Cloud Enabled Small Cells in the context of the 5G ESSENCE project. The data collected by the monitoring system can be valuable for the extraction of knowledge related to user habits, user mobility patterns, spatio/temporal user resource demands, etc. Some examples are presented in [10] that evaluates user mobility models in a University campus, in [11] that makes use of Wi-Fi measurements to analyze mobility patterns in a Hospital or in [12] that assesses the spatio/temporal traffic correlation of Wi-Fi traffic data measurements and presents a supervised learning approach to classify the different campus spaces.

Within this context, this paper presents a new self-planning methodology to determine adequate relocations of APs to improve the provided coverage/quality at those regions of the network with low signal quality and high traffic demands. The methodology makes use of a monitoring system that collects historical network measurements and an analytics system that extracts valuable knowledge from the collected data. This knowledge feeds the AP relocation strategy that is based on a genetic algorithm. The performance of the proposed relocation actions are validated in a real Wi-Fi network. Although the AP relocation methodology considered in this paper focuses on a Wi-Fi network, it could be adapted as well to other contexts, for example exploiting the monitoring and analytics framework for small cell networks defined by the 5G ESSENCE project in [9].

The remaining of the paper is organized as follows. Section 2 presents the different steps of the proposed self-planning methodology, while the details of the AP relocation algorithm are presented in Section 3. The performance of the proposed approach is empirically evaluated in Section 4, while Section 5 summarizes the conclusions.

2 Proposed Self-planning Methodology

The proposed self-planning methodology assumes a Wi-Fi network with monitoring capabilities enabling the collection of a number of measurements reported by the users when connected to the different APs. The collected data is processed by a data analytics module to generate actionable insights to be used by the self-planning process. As a result, the self-planning process will be able to detect those situations in which the network performance is not satisfactory thus requiring relocation of one or several APs to other positions. The network is represented as a group of K Access Points located at specific geographical locations ($x_{AP,k}, y_{AP,k}$). The complete scenario is represented with a set of pixels R, each one associated to a geographical location. According to Fig. 1, the *Network Monitoring* module collects and stores user measurements during a large period of time T_{meas} (i.e. weeks or even months). Each measurement is associated to a particular location s (with $s=1,...,S$) with coordinates (x_s, y_s), i.e. one of the pixels of the scenario, so that $S \subseteq R$ is the set of locations in the scenario with associated measurements. The location information can be obtained by means of different positioning methodologies which are out of the scope of this work, such as Time of Arrival [13] or fingerprint-based approaches [14]. With these inputs, the *Data Analytics* module generates the following metrics for each location:

1. List of APs where the user was connected to as best server at the s-th location.
2. Average Received Signal Strength Indicator ($RSSI_s$) from the best AP at the s-th location.
3. Average Signal to Noise Ratio (SNR_s) at the s-th location.
4. Normalized traffic volume (V_s) defined as the percentage of traffic transmitted/received at the s-th location with respect to the total amount of traffic volume transmitted/received in all the scenario in the period T_{meas}.
5. The activity factor (T_s) defined as the amount of time in which a user at the s-th location is connected to the network with respect to the measurement period T_{meas}.

These historical measurements are very valuable for the self-planning process in order to decide adequate actions of AP relocation. For example, AP relocations may decide to enhance the performance in certain regions with high traffic demands that experience poor SNR or bit rate by moving APs from regions with lower traffic demands (e.g. stairs and corridors of the building where users do not stay connected too much time and do not transmit large traffic volumes).

The *Data Analytics* module also identifies the regions with relatively high traffic demands by selecting the set $U \subseteq S$ where the normalized traffic volume V_s and activity factor T_s are higher than some specific thresholds Th_v and Th_t, respectively. Then, with the set of samples U, a filtering process is carried out to determine the set of samples $W \subseteq U$ in which the provided coverage and quality are below some specific thresholds Th_RSSI and Th_SNR, respectively. Therefore, the set W represents the most relevant locations of the network where users transmit a high amount of traffic and the network coverage/quality requirements are not guaranteed. The set W is used to trigger the *AP relocation process*. Specifically, if the set W is not empty (i.e. there are some relevant locations with poor coverage/quality), the *AP relocation algorithm* described in Section 3 is triggered. Otherwise, if the set W is empty, (i.e. there are no

samples with poor coverage/quality) no AP relocation is done and a new period of time T_{meas} starts.

An adequate setting of thresholds Th_v and Th_t is prime important in order to adequately determine the relevant sample locations considered for optimization. A too low value for these thresholds Th_v and Th_t may cause the appearance of false alarms (i.e. wrong activations of the *AP relocation algorithm*) since the existence of non-relevant locations (i.e. low traffic volume and low activity) with bad RSSI and SNR will unnecessarily activate the *AP Relocation algorithm*. On the contrary, a too high value in these thresholds Th_v and Th_t may cause that the *AP Relocation algorithm* may not be activated when needed. Note also that, in case of activation of the *AP Relocation algorithm*, high values of Th_v and Th_t may cause the methodology to miss relevant sample locations (with relatively high traffic volume and activity), which will lead to a sub-optimal AP relocation. As for the thresholds Th_RSSI and Th_SNR, they must be selected in accordance with the RSSI and SNR values that ensure a certain user performance (e.g. a minimum required user bit rate).

By observing the set of measurements W, the *AP Relocation algorithm* will determine a group of M Access Points ($M \leq K$) considered as candidates to be relocated. These M APs are selected as the best serving APs for the samples of the set W. Then, the objective of the *AP Relocation algorithm* is to move one or several of these APs closer to the locations with high traffic demands (i.e. set of samples U) but, at the same time, maintaining enough coverage at all the R pixels of the scenario. Since the APs are usually connected to the wired network and hanged on the walls, the selection of potential AP locations depends on how easily they can be placed. In this respect, an adequate filtering of unfeasible locations will determine the subset $R_C \subseteq R$ that includes only the feasible locations where the APs can be placed. This filtering will reduce the computational complexity of the solution search, which consists in running a metaheuristic algorithm that iteratively proposes and evaluates different candidate solutions until a termination condition is fulfilled.
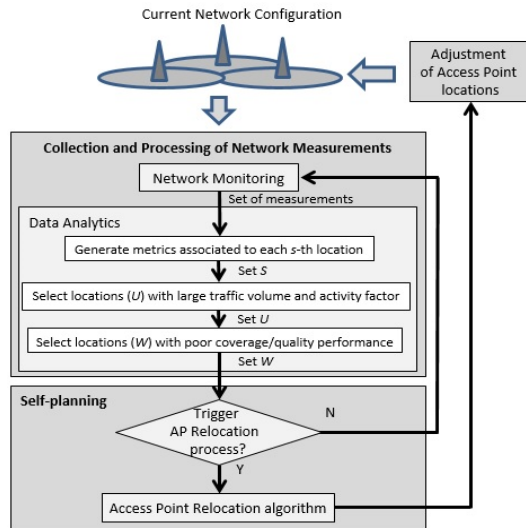

Fig. 1. Self-optimisation methodology.

Different metaheuristic algorithms can be considered (e.g. genetic algorithm [15], simulated annealing [16], etc.). This paper makes use of a genetic algorithm approach which is described in Section 3. If the *AP Relocation algorithm* provides a feasible solution, then, the corresponding AP(s) are relocated to the identified position(s).

3   AP Relocation Algorithm

The *AP Relocation Algorithm* makes use of a genetic algorithm optimization that consists on iteratively proposing possible candidate solutions (*individuals*), each one consisting in a combination of AP locations, and evaluate the performance of each individual according to a cost function. The best individuals in an iteration (*generation*) are combined to obtain new individuals that are again evaluated in the subsequent iteration. This process is repeated during multiple generations G. Each individual is represented by a combination of the geographical coordinates of all the K APs in the network. During the execution of the genetic algorithm, only the M APs ($M \leq K$) that are the best serving APs for the samples of the set W are candidates to be relocated, while the rest of APs will remain at the same position. To following steps summarize the operation of the genetic algorithm for relocating the M APs:

1) *Generate a population of N individuals:* Each n-th individual ($n=1,\ldots,N$) is represented by a vector $\boldsymbol{v_n}$ with the geographical coordinates of the K APs $\boldsymbol{v_n}=\{x_{AP,1}, y_{AP,1},\ldots, x_{AP,k}, y_{AP,k}\ldots, x_{AP,K}, y_{AP,K}\}$. Each element of the vector $\boldsymbol{v_n}$ is called a *gene*. For each individual, the coordinates of the M candidate APs to be relocated are determined randomly with uniform distribution among the subset of feasible locations R_C, while the coordinates of the rest (K-M) APs are fixed at their current locations in the real network.

2) *Execute the following sub-steps for each n-th individual:*

2.1- Determine the path loss L_u between each sample $u \in U$ and its best AP, i.e. the AP with the lowest path loss between the position of sample u and the positions of the APs in the n-th individual according to a predefined propagation model.

2.2- The cost of the n-th individual C_n is calculated as the weighted average of the propagation loss L_u from each u-th sample to its best AP as follows:

$$C_n = \frac{1}{U}\sum_{u=1}^{U} w_u \cdot L_u \tag{1}$$

where the weights w_u are defined to give more relevance to the regions with high traffic volume and low SNR. For this reason, the weights w_u are defined as:

$$w_u = \frac{\alpha_u \cdot \beta_u}{\sum_{u'=1}^{U}(\alpha_{u'} \cdot \beta_{u'})} \tag{2}$$

where:

$$\alpha_u = \frac{V_u}{\sum_{u'=1}^{U} V_{u'}} \tag{3}$$

$$\beta_u = \frac{1/SNR_u}{\sum_{u'=1}^{U} 1/SNR_{u'}} \tag{4}$$

V_u denotes the normalized traffic volume demanded at the *u*-th sample and SNR_u is the Signal to Noise Ratio of sample *u*-th. α_u and β_u are normalized parameters so that $0 \leq w_u \leq 1$ as shown in (3) and (4).

According to equation (1), an individual with a lower cost C_n corresponds to a better network layout since it relocates the different APs so that lower values of path loss are observed, especially for samples with high traffic volume or low SNR.

2.3- Two feasibility conditions are checked. The first condition checks whether the individual guarantees the overall network coverage or not. Then, an individual is not feasible if there is any geographical location in the scenario (i.e. a pixel of the set *R*) with RSSI lower than $RSSI_{cov}$. The second feasibility condition checks that the proposed individual guarantees that the interference generated from an AP to a neighbor AP that is using the same channel is lower than a threshold Th_RSSI_{neigh}.

3) *Selection:* This step determines which feasible individuals are selected for the generation of new individuals to be evaluated in the subsequent generation. The cost of each individual C_n is used for the selection process by considering a *roulette wheel* [16] so that the probability of selecting the *n*-th individual for recombination is:

$$Prob_n = \frac{\frac{1}{C_n}}{\Sigma_{i=1}^{N} \frac{1}{C_i}} \qquad (5)$$

According to this, individuals with lower cost have higher probability to be selected for recombination. Individuals that are not feasible according to the feasibility conditions in step 2.3, are not considered in this selection process.

4) *Recombination:* This process combines the different genes (i.e. elements of the vectors $\boldsymbol{v_n}$) of the two individuals selected in the previous step (called *parents*) to generate two new individuals (called *children*). The rationality of this process is to search for new solutions similar to the best individuals of the previous generation by combining their genes. The recombination process considered here is the so-called "1-point crossover" [16]. Considering the genes of the two parents, a crossover point is defined randomly and all the genes beyond this crossover point are swapped between both parents to obtain the children.

5) *Mutation:* This operator makes small random changes in the genes of the two individuals obtained after recombination. The probability of selecting a gene for mutation is $1/N_{genes}$ where $N_{genes}=2M$ corresponds to the two coordinates of all the *M* AP that can be relocated in the individual. As a result, very few genes of an individual are usually modified by this process. When a gene is selected for mutation, the new value, which represents a new AP location, is either an increase or decrease (with equal probability) in one pixel, allowing only changes to locations of the set R_C.

As a result of the selection, recombination and mutation process, a total of two new individuals will be obtained. This process is repeated *N/2* times until getting the *N* individuals for the new generation. With the newly generated *N* individuals, the algorithm executes again the evaluation procedure and computes the associated costs. The process is repeated iteratively until reaching a maximum number of generations *G*. Then, the final solution of the algorithm is the individual with minimum cost that has been found throughout all the generations.

4 Performance Evaluation

The proposed methodology has been evaluated empirically in a scenario consisting of a building of 40m x 20m covered by $K=3$ APs represented in Fig. 2. A grid with pixel resolution of 1m is considered. The measurements are made and reported by 241 different users during a period T_{meas} of one month. These reported measurements are collected by the Cisco Prime Infrastructure tool [17]. The locations of the set of $U=12$ samples in which the demand of traffic is relatively high (i.e. Normalized Traffic Volume higher than $Th\_v=1\%$ and Activity Factor higher than $Th\_t=1\%$) are represented in Fig. 2. Table 1 presents the average SNR_u, the normalized traffic volume V_u, the normalized parameters α_u and β_u and the weight w_u for the identified $U=12$ samples. By filtering the U samples, the set W contains one sample with RSSI and SNR below $Th\_RSSI=-81dBm$ and $Th\_SNR=15dB$, respectively. This sample corresponds to u_{12} located at coordinates $[39,2]$ as shown in Fig. 2. Although it is just a single sample, it represents a substantial amount of traffic (i.e. 4.83% of normalized traffic volume) and exhibits a very low SNR compared to the rest of samples. At this location, an average bit rate of 14 Mb/s has been observed while the bit rate at other locations inside the building usually varies between 30 and 50 Mb/s. At this location, the users can connect to two different APs, namely AP1 and AP3. Then, the *AP relocation algorithm* will evaluate the relocation of these $M=2$ APs to improve the per-

Fig. 2. Considered scenario with the set of $U=12$ samples.

Table 1.- Observed metrics and obtained weight for each sample of the set $U=12$.

Sample	Location	$SNR_u(dB)$	$V_u(\%)$	α_u	β_u	w_u
u_1	[14,7]	33.04	1.83	0.006	0.050	0.004
u_2	[22,3]	30.12	1.75	0.012	0.048	0.007
u_3	[7,4]	17.54	1.68	0.231	0.046	0.118
u_4	[3,14]	27.24	8.27	0.024	0.227	0.062
u_5	[1,2]	23.95	1.77	0.052	0.048	0.028
u_6	[13,15]	29.32	3.56	0.015	0.098	0.017
u_7	[13,5]	27.13	1.68	0.025	0.046	0.013
u_8	[35,18]	23.82	1.8	0.054	0.049	0.030
u_9	[39,18]	22.80	2.34	0.068	0.064	0.049
u_{10}	[39,14]	22.08	1.78	0.081	0.049	0.044
u_{11}	[37,18]	33.70	4.96	0.005	0.136	0.008
u_{12}	**[39,2]**	**14.92**	**4.83**	**0.427**	**0.139**	**0.620**

formance at the $U=12$ locations, while maintaining the network coverage (i.e. RSSI of all the R pixels in the scenario is higher than $RSSI_{cov}=-75dBm$) and keeping the interference between neighbor APs using the same channel below $RSSI_{neigh}=-70dBm$.

The genetic algorithm considers a population of $N=50$ individuals, each one represented by the location coordinates of the $K=3$ APs. A total of $G=5000$ generations have been considered. The propagation model is taken from the ITU-R [18] so that the path loss between the transmitter and the receiver is calculated as:

$$L(dB) = 20log_{10}(f) + N_o log_{10}(d) + P_f(n) - 28 + N_d L_o \qquad (6)$$

where $f=2.4GHz$ is the carrier frequency, $N_o=30$ for office areas, $P_f(n)$ is the floor attenuation factor considered here 0 dB since all the APs and the users are located at the same floor, N_d is the number of obstacles (i.e. walls and doors) between the transmitter and the receiver and $L_o=2dB$ is the path loss associated to each obstacle. After executing the *AP relocation algorithm*, the best solution found is illustrated in Fig. 3. It consists in moving AP3 to the location *[32,2]* while keeping AP1 in its current location. This solution has a cost of 75.58 which is considerably better than the cost of 93.6 that was obtained before the execution of the *AP relocation algorithm*.
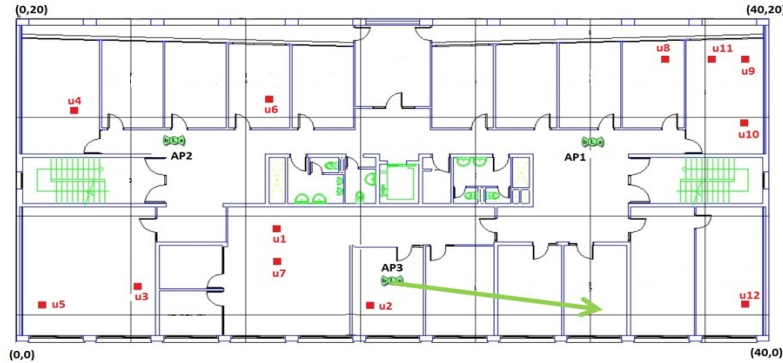

Fig. 3. Solution proposed by the Genetic Algorithm.

In order to empirically validate the solution proposed by the genetic *AP relocation algorithm*, several RSSI measurements have been done at 47 different locations of the building before and after the AP relocation (i.e. with AP3 located at coordinates *[20,4]* and *[32,2]*, respectively). For each location, the RSSI averaged in a period of 2 minutes is obtained. Fig. 4 and Fig. 5 present the average RSSI measurements obtained before and after the AP relocation, respectively. After the AP3 relocation, an average $RSSI=-67dBm$ has been measured at u_{12} (i.e. coordinates *[39,2]*) which is considerably better than an average $RSSI=-82dBm$ at the same location before the AP3 relocation). Moreover, the AP relocation guarantees an RSSI higher than the threshold $RSSI_{cov}=-75dBm$ for all the locations of the scenario. Some other validations have been done by moving AP3 to different locations and collecting the new values of the average RSSI measurements. For example, moving AP3 closer to the location of

sample u_{12} (e.g. to coordinates *[37,2]*) is not a feasible solution since it causes some coverage problems at region around *[17,8]*. Similarly, relocating AP3 to coordinates *[28,2]* (i.e. not so close to u_{12}]) guarantees the coverage RSSI requirements at all the locations of the building but, an average RSSI=-73dBm is measured at u_{12}, which is not as good as the average RSSI=-67dBm obtained when implementing the best solution proposed by the genetic *AP relocation algorithm*.

Fig. 4. RSSI measurements obtained before the AP Relocation i.e. AP3 at *[20,4]*.

Fig. 5. RSSI measurements obtained after the AP Relocation i.e. AP3 at *[32,2]*.

5 Conclusions

This paper has presented a new self-planning methodology to determine adequate relocations of APs in a Wi-Fi network based on information about signal quality and user traffic demands extracted from network measurements collected and processed by a monitoring and analytics system. From these measurements, the proposed methodology identifies regions with poor RSSI/SNR and high traffic demand, determines candidate APs to be relocated and applies a genetic algorithm optimization to recommend new locations for these APs. The paper has presented an empirical validation of the algorithm performance, making use of measurements collected from a real Wi-Fi

network and comparing the obtained RSSI at different locations before and after the execution of the AP relocation.

References

1. Cisco Visual Networking Index, "Forecast and Trends, 2017-2022", Cisco white papers November, 2018.
2. J. Ramiro, K. Hamied, *Self-Organising Networks. Self-planning, self-optimization and self-healing for GSM, UMTS and LTE*, John Wiley & Sons, 2012.
3. H. Gacanin, A. Ligata, "Wi-Fi Self-Organizing Networks: SON challenges and use cases", *IEEE Communications Magazine*, Vol. 55 issue 7, 2017.
4. C. Singhal, D. Swades, "eM-SON: efficient multimedia service over self-organising Wi-Fi network", *IET Networks*, January 2018.
5. P. Muñoz, O. Sallent, J. Pérez-Romero, "Self-Dimensioning and Planning of Small Cell Capacity in Multi-Tenant 5G Networks", *IEEE Trans. on Vehicular Tech.*, May, 2018.
6. A. Politi, L. C. Zamboni, R. Pasti, L. N. de Castro, "An optimization model for the indoor access point placement problem with different types of obstacles", *Latin America Congress on Computational Intelligence (LA-CCI)*, 2015.
7. R. Atawia, H. Gacanin, "Self-deployment of future indoor Wi-Fi networks: An artificial intelligence approach", *IEEE Global Communications Conference, (GLOBECOM)*, 2017.
8. I. C-L, Y. Liu, S. Han, S. Wang, G. Liu, "On Big Data Analytics for Greener and Softer RAN", *IEEE Access*, August, 2015.
9. J. Perez-Romero, V, Riccobene, F. Schmidt, O. Sallent, E. Jimeno, J. Fernández, A. Flizikowski, I. Giannoulakis, E. Kafetzakis, "Monitoring and Analytics for the Optimisation of Cloud Enabled Small Cells", *IEEE Intl. Workshop on Computer-Aided Modeling Analysis and Design of Communication Links and Networks (CAMAD)*, September, 2018.
10. J. Kim and A. Helmy, "Analysing the mobility, predictability and evolution of wlan users," *Intl. Journal of Autonomous and Adaptive Communications Systems*, 2014.
11. A. Ruiz-Ruiz, H. Blunck, T. Prentow, A. Stisen, and M. Kjaergaard, "Analysis methods for extracting knowledge from large-scale wifi monitoring to inform building facility planning", *IEEE Internat. Conf. on Pervasive Computing and Communications*, March 2014.
12. A. E. Redondi, M. Cesana, D. M. Weibel, E. Fitzgerald, "Understanding the Wifi usage of University Students", *Internat. Wireless Comm. and Mobile Computing Conference*, 2016.
13. M. Ciurana, F. Barceló, "Indoor tracking in WLAN location with TOA measurements", *4th Int. Workshop on Mobility Manag. and Wireless Access*, Torremolinos (Spain), 2006.
14. X. Wang, L. Gao, S. Mao, S. Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach", *IEEE Trans. on Vehicular Technology*, January 2017.
15. E. Aarts, J. Korst, *Simulated Annealing and Boltzmann Machines*, J. Wiley & Sons, 1989.
16. L. Davis (ed.), *Handbook of genetic algorithms*, Internat. Thomsom Computer Press, 1996.
17. *Cisco Prime Infrastructure 3.5 Administrator Guide*, available at www.cisco.com
18. Propagation data and prediction methods for the planning of indoor radio communication systems and the radio local area networks in the frequency range 900 MHz to 100 GHz, ITU-R Recommendations, Geneva, 2001.