Designing Resource-on-Demand Strategies for Dense WLANs

Designing Resource-on-Demand Strategies for Dense WLANs

Fikru Getachew Debele, Michela Meo, Daniela Renga, Marco Ricca, and Yi Zhang

***Abstract*—Being cheap and easy to deploy, dense WLANs are becoming the most popular solution to provide Internet access in locations in which the population of users is large, as in Campuses, large enterprises, etc. The large density of access points (APs) comes from the need to have enough capacity to carry the traffic generated at peak hours although, in these scenarios, traffic varies a lot on a daily, weekly or seasonal basis. During low or no traffic periods, APs are underutilized, even if they are consuming energy almost in the same amount as if they were fully loaded. Promising solutions to reduce this form of energy waste consist in activating only the number of APs that is strictly needed to carry the actual traffic; in other words, to make capacity dynamically adaptive through Resource on-Demand (RoD) strategies. In this paper, we investigate the case of a portion of the dense WLAN in our Campus. Through real trace analysis, we investigate users behavior in accessing the WLAN and formulate a stochastic characterization of it. We propose a simple model that describes RoD strategies and use it to study the system performance that is evaluated in terms of AP activity and inactivity periods, AP switching frequency, and energy saving. Finally, we present some results obtained by experimenting RoD strategies in a portion of the WLAN. Our results show that RoD strategies for dense WLANs are feasible and effective in trading-off the opposite needs to save some energy and to guarantee a smooth network operation and high quality of service.**

***Index Terms*—Resource-on-Demand, dense WLANs, energy saving, modeling, trace analysis.**

I. Introduction

Wireless local area networks (WLANs) have become one of the most popular solutions to provide reliable, portable and high-speed Internet connectivity to end users [1]. In particular, *Dense WLANs*, i.e., featuring thousands of access points (APs) per square kilometer, are usually deployed where, as in public locations or in large enterprises, a lot of users require WiFi access to the Internet. In these locations, in order to provide enough capacity for the large amount of traffic generated by the users, the density of APs is high, much higher than what needed for coverage only. However, depending on the people behavior and habits, this large amount of capacity is not needed all the time; typically there are long periods of time during a day, or even whole days in a week, in which there is not so much Internet traffic and APs become idle. The power consumption of an idle AP is nearly the same as at full load; thus, during low or no traffic periods, the energy consumed by the AP is "wasted". Energy can be saved by applying an AP switching on/off policy, so that only the adequate number of APs needed for serving the actual traffic is activated. In an era of dramatic concerns about sustainability of communication networks [2], we need to change from an always on communication infrastructure to a *Resource on-Demand* (RoD) paradigm [3]. In a similar way as what we normally do with lighting (we switch on the light when entering a room and switch it off when we leave), we must get used to the idea that the communication infrastructure is activated only when needed.

In a dense WLAN, the concept of RoD means that APs dynamically switch on and off based on users' need for capacity. The strategies to decide AP switching have to trade-off among opposite needs. Energy can be saved by activating a small number of APs, but an adequate number of APs have to be activated so as to guarantee proper levels of quality of service, i.e., bandwidth per connection. Moreover, frequent AP switching should be avoided to guarantee a smooth network operation. In this paper, we focus on the design, analysis and parameter setting of RoD strategies for dense WLANs, investigating these trade-offs. In particular, we consider a dense WLAN that provides WiFi Internet access in our Campus; Politecnico di Torino (PoliTo), in which more than 300 APs have been installed. These APs are powered by Power over Ethernet (PoE) switches and managed by a proprietary central controller that is able to configure, monitor and collect statistics over the network utilization. As specified in the AP manufacturer data sheet, a single device draws a maximum power of 13 W [4]. In our experimental activity, we observed a nominal power consumption of about 8 W. Using this latter figure, at Campus wide the power consumption contribution related to the dense WLAN corresponds to a total of 2.4 kW. Furthermore APs are active 24 hours over 7 days which implies that the estimated consumption over the whole year is about 21 MWh. While the monetary cost of running the WLAN is relatively low, in an era of sustainability concerns, it is desirable that any energy waste is avoided. Reduction of energy wastes reduce useless costs and has also the beneficial effect of contributing to change people habit towards a more parsimonious and aware use of energy.

In our work, we collected real traces, we analyze users' behavior in accessing the network, we design RoD strategies by modeling the system and we test them in a portion of the operative network. Our experimental setup, from which we collect traces and test RoD strategies, is established in a study room inside the Campus and was already presented in our previous work [5]. More in detail, our contribution is organized in the following steps:

- First, from real traces, we investigate the behavior of the

Fikru Getachew Debele, Michela Meo, Daniela Renga, Marco Ricca and Yi Zhang are with the Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Torino, TO, 10129 Italy e-mail: name.lastname@polito.it.

users in accessing the WLAN. We find that during the day, except for transient periods early in the morning and late in the afternoon, the user arrival rate is quite constant and the interarrival time follows quite well an exponential distribution. The distribution of session duration is more complex, but to an acceptable level can be approximated by an exponential distribution also.
- Second, we model the users' behavior in the system as an $M/M/\infty$ queuing system and validate this model by comparing theoretical results against observations of the real system. The comparison shows a good match between the theoretical model and measurement results.
- Third, based on the inferred model mentioned above, we build a Markov chain in which the AP activation and deactivation mechanisms are described. We solve the model deriving a number of interesting metrics, such as the AP activation time, the duration of the AP idle periods, and the switching frequency. A deep investigation of these metrics allows us to understand the performance of the RoD strategies and to derive some indications on the parameter setting and design of the strategies.
- Finally, we build experiments on our testbed and compute the energy savings; we estimate the potential saving that can be achieved with a wider deployment of RoD strategies in the whole Campus WLAN.

This paper is organized as follows. Section II discusses some related work. Section III presents the considered system, while Section IV investigates WiFi users' behavior and demonstrate the fitting of the distribution of interarrival time and session duration with exponential distributions. Models to describe the system behavior and to define parameters setting for RoD strategies in dense WLANs is presented in Section V. RoD strategies' performance is evaluated in Section VI, the model is used to investigate the impact of the RoD parameters on energy saving, system stability and the quality of service. As a proof of concept of the operation of RoD strategy, we present experimental results in Section VII. Finally we conclude the paper in Section VIII.

II. Related Work

In this paper, we focus on both system analysis through traffic measurement and energy saving through resource on-demand strategies. In this short discussion on the related work, we therefore separately discuss the literature that is mostly related to our work in these two fields.

Modeling the WiFi users' behavior, typically based on traces analysis collected from WLANs, has been studied for more than 10 years since WLAN was beginning to be widely deployed all over the world. However the large deployment of APs was proposed from the constraint to provide bandwidth, especially during peaks hours, to users. This direction brings about the presence of a significant energy consumption and suggest to investigate on energy-efficiency WLAN, that has become a hot topic in the recent years. To better present our work we focus on both system analysis through users' characterization over the dense WLAN and the design of a strategy able to adapt dynamically the AP switching on/off in order to save energy. To the best of our knowledge, our work address these two aspects together. Indeed, in [6] authors investigated users' characteristics in different public locations, e.g., library in universities, cafeterias on the street, etc, showing fittings of probability distributions of packets and flows. Besides, authors of [7] proposed a Poisson-arrival model and a Markov chain to analyze still the TCP and UDP traffic in WLAN, for resource management by dynamic spectrum access. In [8] the measurement of different types of traffic, e.g. HTTP, P2P, SSH, etc., was investigated in public WLAN hotspots. In [9] and [10] authors measured users characteristics in the Campus WLAN of University of North Carolina, fitting them with different distributions such as Poisson, BiPareto and Lognormal. The authors in [11] analyzed the measured traffic in WLAN, trying to eliminate redundant traffic to improve the performance of the network. These related works give us hints of how to measure, analyze, and model users behavior in accessing the Campus dense WLANs. Furthermore, these works did not consider the possibility of using the measurement data for the purpose of energy saving, which has become an important issue to take into account today.

For what concerns energy efficiency, our approach is based on the RoD concept that has already been proposed for the wireless access, in particular for cellular networks. In the field of cellular networks, the work [12] introduced the idea of switching on and off the Base Stations (BSs) of a mobile network by scheduling their activity based on historical data of traffic, while in other works, such as [13], [14], BSs switching is decided based on real-time traffic measurements. The idea is similar to what we are considering here, but the implementation, as well as the constraints are quite different: in our case, dense WLANs make use of a centralized controller that can quite effectively take decisions, in the case of cellular networks a particular care should be taken in the coordination among BS decisions [15], [16]. For a survey on RoD strategies for cellular and WiFi networks see also the survey in [17]. In the context of WiFi networks, among the first work to specifically consider RoD in dense WLANs is [3]. Most of the existing works related to energy-efficient WLAN are based on theoretical and/or simulation analysis without real experiments, e.g., [18]–[20], while we measure the real-time user traffic and behaviors and take real-time decisions about AP switching on/off based on measured data. There is only a few work on implementing energy-efficient WLAN testbeds. For instance, authors of [21] implemented an experimental system to switch on/off WLAN APs by using wireless sensors to detect new incoming users, and in [22] authors propose a prototype using a station to wake on the WLAN when all the APs are in the sleep mode. Our work is based on a previous work [23] proposed by authors that belong to our academic research group that proposed thresholds and hysteresis windows for switching on/off APs in dense WLANs. Moreover we deployed a RoD testbed in PoliTo Campus production network, the outcome of that work was described in [5], and all the measurement and experiment results in this paper are obtained from this testbed. The testbed is located in a dense WLAN where there are multiple APs covering the same area so that one AP is always on for guaranteeing

the coverage and detecting incoming users. Therefore, no additional wireless sensors or devices are needed. Besides, our testbed works with any kind of commercial APs that support PoE without requiring any software and hardware modifications inside the APs firmware.

III. System Description

Enterprise as well as educational institution WiFi networks are characterized by *dense* deploymentof APs to provide high-capacity connectivity as well as wide coverage. The WiFi network of PoliTo is one of such deployments: in highly dense locations of our Campus, such as classrooms and libraries, the AP density is as high as $0.02[AP/m^2]$; by comparison, consider that in lightly loaded environments an AP can provide a coverage of $2500m^2$ [24], [25], corresponding to a density of about $0.0004[AP/m^2]$, two orders of magnitude smaller than in our case.

In this research activity, we focus on a study room equipped with 3 APs that provide connectivity to up to 120 – 140 students at the same time. The needed capacity is high and the connectivity provided by the 3 APs is fully overlapping. The considered study room, similarly to other areas of the PoliTo Campus WLAN, such as the library, and meeting venues, exhibits particularly high demand for connectivity. The methodology applied here is quite general and can be used for studying other areas of the Campus WLAN with less demand (for example, classrooms and offices). To analyze the system under study, we collect traces from the 3 APs installed in the study room. Traces are collected from a central control system that manages the whole PoliTo WiFi network. Users need to be authenticated with the control system to use the WiFi network. Once a user is authenticated, the controller creates and saves a session record corresponding to that particular user. The session ends when the user logs out of the system or the WiFi network is unreachable. A user session is represented as a row in the trace with more than 20 fields of parameters. In this study, for the purpose of system characterization, we are only interested in the following fields:

- Association time
- Disassociation time
- Session duration

These data allow us to investigate the way in which users access the APs, that is fundamental for the design of RoD strategies. Users' behavior and WiFi usage change on both a daily basis as well as on longer time scales. To have a deeper understanding, we focus on two different periods, that are typical of the academic activity: a normal *teaching* period, with lecturing activity and many students often coming in and going out of the study room; and an *exam* period, with slightly fewer users in the room. In particular, the trace analysis in this work includes the following periods:

- **Teaching period**: traces are collected between October 27th, 2014 and December 19th, 2014
- **Exam period**: traces are collected between January 21st, 2015 and February 24th, 2015

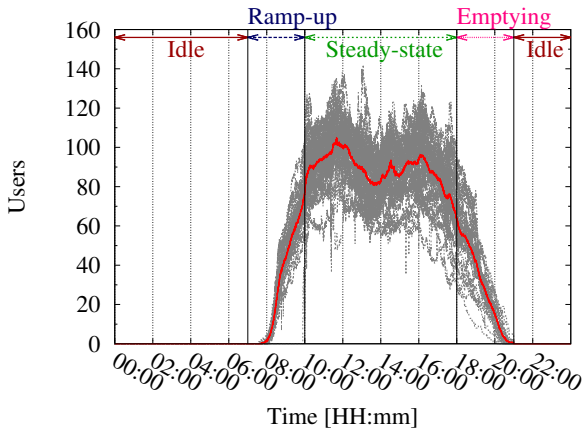


Fig. 1: Number of users in the system for the days of the teaching period (Mean value in red)

IV. Trace analysis

In this section, we investigate the behavior of the users in the system under study through the traces collected in the periods identified in previous section. Fig. 1 shows the daily number of users associated to the three APs for working days of the teaching period. Each gray curve corresponds to a day and the solid red line represents the average number of users in the system; each sample was measured on time intervals of one second. The daily patterns are extremely repetitive, curves are very similar to each other. In general, four phases can be identified in a day:

- **Idle** phase, between 21:00 to 7:00 of the day after. No users is in the room, actually the Campus is closed.
- **Ramp-up** transient phase, between 7:00 to 10:00. Users are arriving in the room at a quite fast rate, the number of users in the rooms grows quickly up to its typical value.
- **Steady-state** phase, between 10:00 and 18:00. The number of users varies around its typical value.
- **Emptying** transient phase, between 18:00 to 21:00. The number of users decreases.

From the trace analysis we extract the two most important parameters to characterize the users' behavior:

- the **interarrival time**, i.e., the time between two consecutive user arrivals at the system;
- the **session duration**, i.e., the time spent by a user associated to an AP.

We focus on these two parameters and characterize them as random variables that can then be used for analytical modeling.

A. Users' interarrival time

The trace analysis is performed at two observation granularities. We start with a general understanding of the system through a daily based trace analysis; then we observe and model the system at a smaller granularity based on single hours.

1) Daily analysis: To understand the evolution of users' arrivals over the whole day, we present in Fig. 2 the total number of users' arrived in the system during a day. Every line

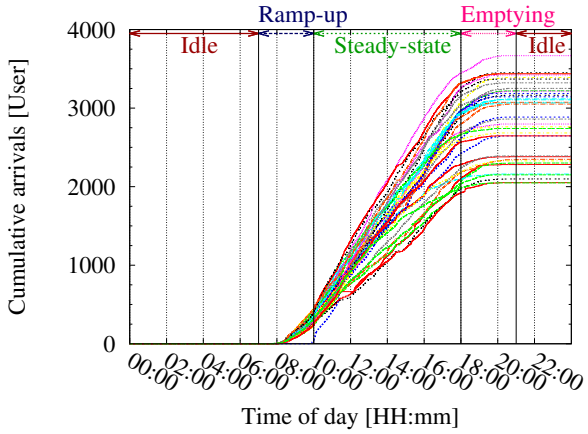


Fig. 2: Daily total number of users arrived in the system for the days of the teaching period.

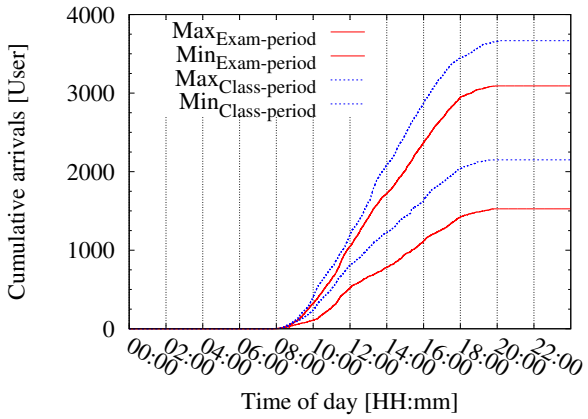


Fig. 4: Total number of arrivals in a day with minimum and maximum number of arrivals, for teaching and exam periods.

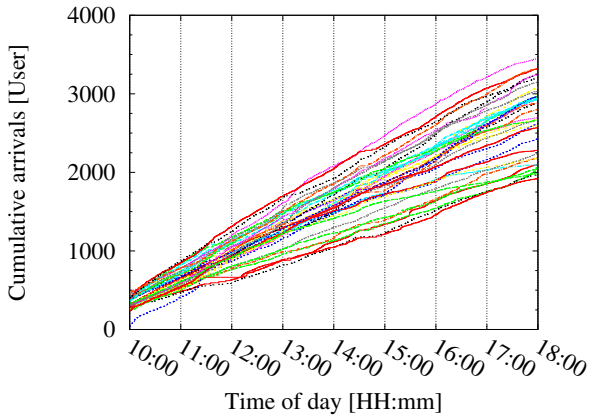


Fig. 3: Total number of arrived users in steady-state phase of the days of the teaching period.

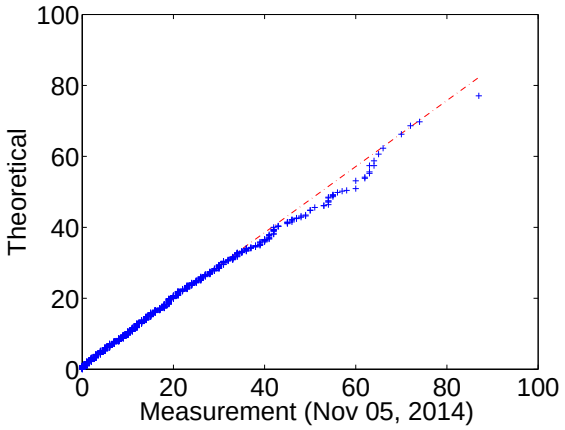


Fig. 5: Q-Q plot of the users' interarrival time; steady-state phase of teaching period.

corresponds to a single day of the teaching period dataset. The figure reveals the same four daily phases that were identified above: idle, ramp-up, steady-state and emptying. Interestingly, during steady-state phase, the growth of users' arrivals follows an almost linear trend denoting a roughly constant arrival rate. Combining this with the observation of Fig. 1, we can say that at the steady-state phase the system is highly dynamic, with users coming and going resulting in a total number of users in the system that varies around a typical value.

To further investigate the steady-state phase, Fig. 3 reports the total number of users arrived in the system for the steady-state phase only. The zoom confirms that while there is quite a different rate in different days, the rate tends to be constant in the same day for the whole steady-state phase. The users' arrival rate during the steady-state phase of the teaching periods varies in the range $[0.0565, 0.1116]$ [user/second] (or $[3.3896, 6.6960]$ [user/minute]).

The behavior during the exam period is similar, but with smaller typical values and smaller arrival rates. To summarize this and, at the same time, compare the exam period with the teaching period, we show in Fig. 4 the number of arrived users for the days that exhibit the maximum and the minimum number of arrived users in both the considered periods. The exam period exhibits the same qualitative behavior of the teaching period (with the same four phases already discussed). The arrival rate is smaller, with a reduction of about 30% with respect to the teaching period. The range of arrival rates in the exam period is in $[0.0412, 0.0943]$ [user/second] (or $[2.4689, 5.6570]$ [user/minute]).

We now go deeper into our investigation and proceed with a fitting of the random variable that represents the users' interarrival time. We focus on the steady-state phase and follow a twofold approach: 1) we study the statistics by curve fitting over the whole observation window for single days, and 2) we repeat the curve fitting, fixing a day, and studying the hourly behavior.

Given the large number of *potential users* arriving at the system and the large variety of different patterns of users' behavior, we formulate the hypothesis that, during the steady-state phase, the arrival process can be represented by a Poisson process. To validate this assumption, we compare the users' interarrival times with instances of a exponentially distributed random variable with mean value equal to the mean interarrival time measured from the traces in an observation window. Fig. 5 reports the Q-Q plot (quantiles of the empirical distribu-

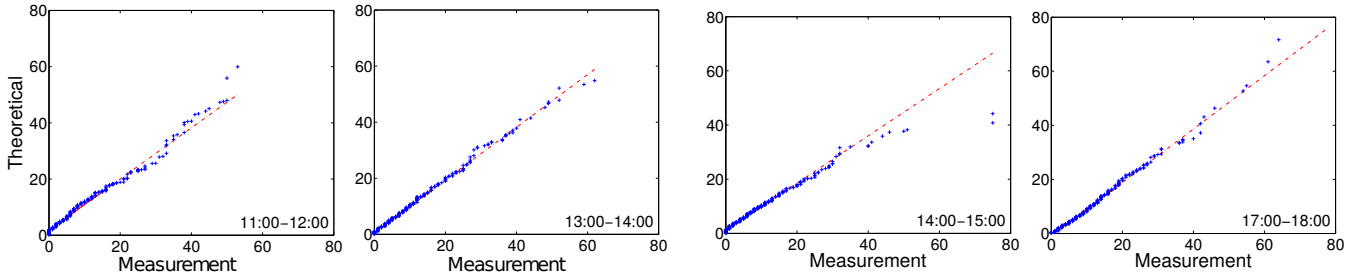


Fig. 7: Q-Q plot of the users' interarrival time; steady-state phase of teaching period, hourly trace analysis.

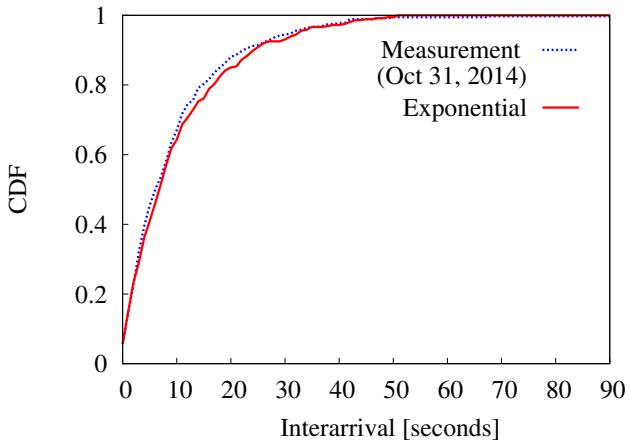


Fig. 6: CDF of users' interarrival time; exponentially distributed theoretical distribution and empirical distribution; teaching period, one day between 10:00 and 11:00.

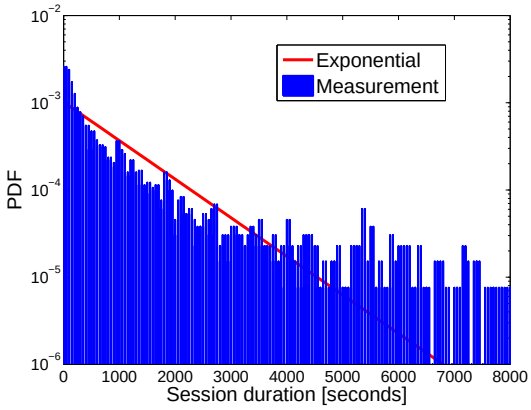


Fig. 8: PDF of the session duration for the empirical data and an exponentially distributed random variable; steady-state phase in the a day of the teaching period (Nov. 3, 2014).

tion versus quantiles of the theoretical exponential distribution) observed in the whole steady-state phase of one day of the teaching period. The figure shows that the exponential distribution is very well fitting the empirical distribution.

2) Hourly analysis: The hourly based trace analysis divides the steady-state phase of a daily trace into hourly traces. Fig. 6 reports the comparison of the theoretical and empirical cumulative distribution function (CDF) of the interarrival time in the hour between 10:00 and 11:00 (Oct 31, 2014), as an example (other hours lead to similar results). The figure validates the assumption that the interarrival time is well fitted by an exponentially distributed random variable even at hourly time granularity.

Fig. 7 shows the Q-Q plot of the empirical data against the the theoretical exponential distribution with parameters obtained from the empirical data in different hours of the steady-state phase of the traces collected on Nov. 4th, 2014. For the sake of clarity some hours are missing in this plot (they are, however, quite similar to plots shown here). The plots confirm that the fitting with the exponential distribution is very good. We also mathematically quantified the claim using a chi-square goodness-of-fit. We test the hypothesis that the interarrival data comes from an exponential distribution. We perform the chi-square test for hourly based traces collected on Oct. 31st, 2014 at a significance level of 0.05 and report the output in Table I. As shown in each column, the measured deviation is always less than the critical value $\chi^2_{0.95,k-1}$ with 95% confidence interval and $k-1$ degree of freedom. Thus, the chi-square test does not reject the null hypothesis for exponential distributed data.

B. Session duration

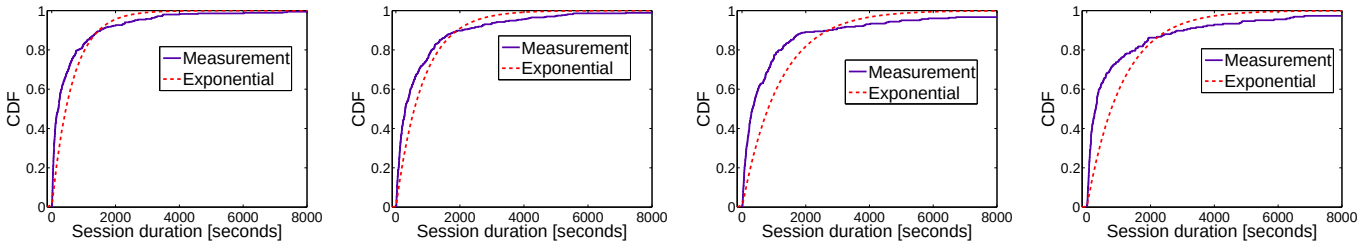
We proceed investigating the characteristics of session durations, the amount of time during which a user is associated to an AP. Again, we start by focusing on the steady-state phase of days in the teaching period and we investigate the distribution both hour by hour and on a daily basis. The average session duration observed over the whole teaching period is equal to 1176 seconds, i.e., about 20 minutes. Table II shows the mean value of the session duration in different hours of the steady-state phase of a single day (we consider Nov. 3rd, 2014) and the mean computed by averaging observations in the whole teaching period, and the whole exam period. The table shows that there is quite a variability in different hours, reflecting the different usage that students make of the room. Sessions are typically longer in the morning and early afternoon, possibly due to the students remaining long time in the study room. Sessions are shorter in the afternoon and at lunch time. This behavior, combined with the constant arrival rate that was previously discussed, translate in the daily pattern of users in the system that is reported in Fig. 1, with a gap at lunch time and a reduction of the number of users from mid afternoon. Clearly, when the average of the whole teaching period is considered, variations are smoothed out. Besides, the exam

TABLE I: Results of the chi-square goodness-of-fit test

	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00	15:00-16:00	16:00-17:00	17:00-18:00
χ^2	3.101	3.686	3.662	2.404	4.143	1.883	4.143	0.621
$\chi^2_{0.95,k-1}$	12.592	7.815	9.488	7.815	9.488	9.488	9.488	5.991

TABLE II: Mean value of the session duration [second] for different hours of a day (Nov. 3, 2014) of the teaching period and for the aggregation of all the data of the teaching and exam periods.

	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00	15:00-16:00	16:00-17:00	17:00-18:00	Daily average
Nov. 3	1094	1069	641	1345	1212	839	833	992	982
teach. period	1043	1198	1093	1277	1356	1088	1007	1014	1176
exam period	1158	950	945	1251	1343	1019	1201	901	1072



(a) 12:00-13:00 (b) 16:00-17:00 (c) 12:00-13:00 (d) 16:00-17:00

Fig. 9: CDF of the session duration for the empirical data and an exponentially distribution random variable: (a), (b) different hours of a day of the teaching period (Nov. 3, 2014): (c), (d) different hours of a day of the exam period (Jan. 27, 2015)

period has an average value of 1072 seconds that is lower than the average session duration of teaching period.

Fig. 8 shows the histogram of session duration in the same day as before (Nov. 3rd, 2014) during the steady-state phase. The histogram is compared with the probability density function (PDF) of an exponentially distributed random variable with the same mean. The plot is represented in a semi-log scale. As expected, the exponential distribution is not fitting that well the empirical data; with a decay that is faster than in the real case.

For the same day, we also focus on the hourly distribution. Fig. 9a-9b show the CDF for both empirical data and an exponentially distributed random variable in two different hours (12:00-13:00, 16:00-17:00) when the mean values of session durations are different (see Table II). Again, the fitting is acceptable but not that accurate. Similarly, Fig. 9c-9d show the CDF for a day in the exam periods (Jan. 27, 2015). The results are slightly different from the results in the day in the teaching period. In the day of the exam period, the users have higher probability to connect to the APs with shorter duration, and the gap between the empirical data and exponential distribution are larger than the day in the teaching period. The distribution for other hours in the day and for other days is quite similar to the case we are showing here. We omit other figures for the sake of brevity.

V. Designing RoD Strategies

In this section, we propose simple analytical modeling techniques that, starting from the analysis of the traces that was previously presented, can be used to design RoD strategies, to understand their performance, and to decide the best setting of the parameters.

A. System Model and Validation

In the previous section, we characterized the system under study focusing on users' interarrival time and session duration. In particular, we observed that users arrive in the target study room according to a Poisson process with rate λ. Our analysis reveals that the value of λ can change day by day but, given a day, λ remains roughly the same over the steady-state observation period that lies between 10:00 and 18:00. The session duration has, as expected, a more complex behavior, with variations both day by day and within hours of the same day. The distribution of the session duration can be approximated by an exponential distribution but the approximation is somehow inaccurate in some cases, typically with heavier tail than exponential. Since the amount of time for which a user accesses the network (i.e., the session duration) can be assumed to be independent on the number of users that are at the same time in the system and since there is no limitation on the maximum number of users, we can model the system as an $M/M/\infty$ queuing system, where the customers represent the students arriving in the room and associating to the WiFi network and the service time represents the duration of a session. Notice that the session duration is not well approximated by an exponential distribution; however, some important results on the $M/M/\infty$ queue, such as the steady-state distribution of the number of users in the system, are independent of the *distribution* of the service time and depend only on the *average* value. Let the average service time, i.e., the average session duration, be denoted by $1/\mu$. From well-known results of queuing theory, the steady-state probability π_i that there are i customers in the queue is given by,

$$\pi_i = \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i e^{-\lambda/\mu} \quad (1)$$

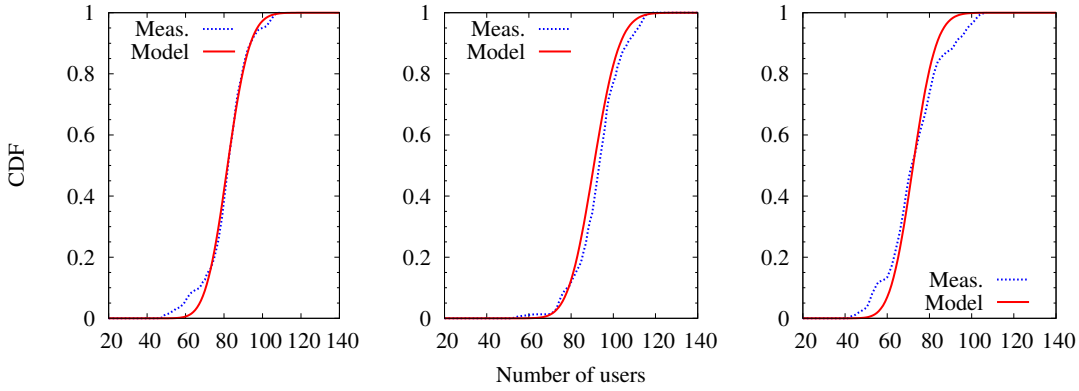


Fig. 10: Comparison between model and actual number of users in the system (27/11 left, 11/11 center, 19/11 right)

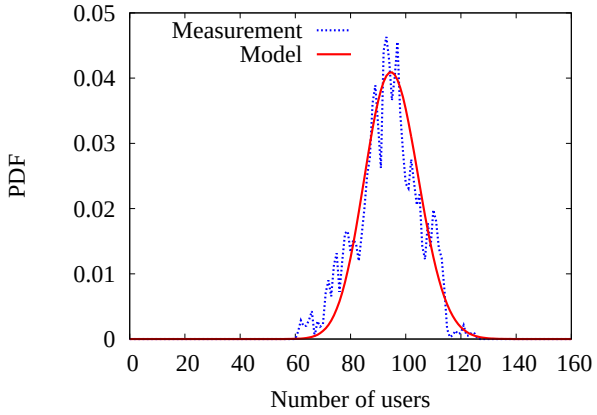


Fig. 11: Comparison between model and actual number of users in the system (10:00 – 11:00)

Before using it to design RoD strategies, we validate here the proposed model. To do this, we compare the number of users observed in the real system with the predictions obtained by the model. As shown in previous section, we focus on both daily and hourly behaviors. For the sake of simplicity, we show only selected results but the conclusions that we derive from the results hold also in, and can be extended to, other days.

As a first step, we observe the steady-state phase of a given day and compute the distribution of the number of users in the room and compare this distribution with the one of the $M/M/\infty$ queuing system, (1). Fig. 10 shows the comparison of the CDF of the number of users as obtained from the measurements and from the analytical model. For all of them, both arrival rate and average session duration are computed over the steady-state period. Observe the good matching between the two curves that confirms the validity of our approach and the accuracy of the model.

To further validate the model, we also extracted and compared the hour by hour users distributions. Fig. 11 shows the PDF for the hour between 10:00 and 11:00. Clearly, in this case, at a finer granularity, the curve obtained from the real traces is noisy when compared with the theoretical model, but, again, the accuracy of the model is very good.

B. Modeling RoD strategies

Now, we use the model presented above to investigate, design and define parameter setting for Resource on-Demand (RoD) strategies in dense WLANs. Our algorithm of RoD is based on the concepts presented in [23]. The capacity demand, and hence the number of active access points (APs), is decided based on the number of users that are associated to APs. This means that when the number of users associated to APs grows above a given threshold, a new AP is activated. A hysteresis loop avoids frequent switching on/off of APs. The basic idea of the hysteresis loop is that the threshold on the number of associated users that we set for deciding to shut down APs is lower than the threshold that we set for deciding to switch on APs. To describe on/off strategies based on hysteresis loop, we assume there is a cluster consisting N APs that provide equivalent coverage in the area of interest. We define the following notation:

- k: number of active APs in a cluster ($k \leq N$)
- $T_h^{(k)}$: *high threshold* for switching on an AP. This threshold is set such that $T_h^{(k)} = kM$, where M is the number of users that an AP can handle with desirable QoS; M is typically specified by the system administrators. When the number of users grows from $kM-1$ to kM, if possible, a new AP is activated.
- $T_l^{(k+1)}$: *low threshold* for switching off an AP.
- ω: *hysteresis width*, such that $T_l^{(k+1)} = T_h^{(k)} - \omega$. An AP is switched off when the number of users decreases to $kM - \omega$.

The switching strategy for turning on/off an AP in a cluster is as follows:

- if the cluster is in k active APs state, an additional AP is activated if the number of users grows beyond the cluster's current high threshold $T_h^{(k)}$;
- if the cluster is in $k+1$ active APs state, an AP is turned off if the number of users goes to the cluster's current low threshold $T_l^{(k+1)} = T_h^{(k)} - \omega$;
- otherwise, retain the number of active APs.

Modeling our system as an $M/M/\infty$ queuing system can be useful to take decisions on the proper setting of some parameters at the basis of the RoD strategy, such as the high and low thresholds, $T_h^{(k)}$ and $T_l^{(k)}$, according to which an

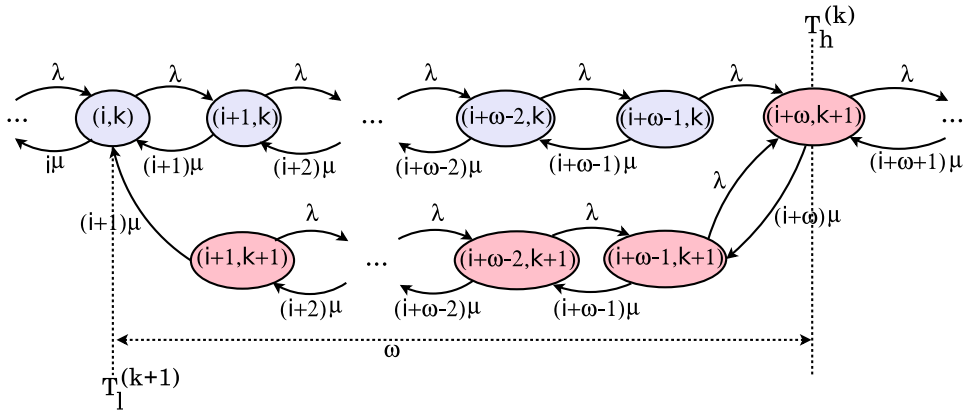


Fig. 12: Markov chain description of hysteresis loop operation

TABLE III: Transition rates from state (i,k) for the MC representing the system with hysteresis window

Destination state	Rate	Condition
$(i+1,k)$	λ	$i \neq T_h^{(k)} \vee k = N$
$(i+1,k+1)$	λ	$i = T_h^{(k)} \wedge k < N$
$(i-1,k)$	$i\mu$	$i \neq T_h^{(k-1)} - w \vee k = 1$
$(i-1,k-1)$	$i\mu$	$i = T_h^{(k-1)} - w \wedge k > 1$

AP is switched on or off, respectively, given that k APs are already active. The size of the hysteresis window between the two thresholds affects the system stability and the energy consumptions in opposite ways: wider values provide higher stability to the system but smaller energy savings, whereas smaller values of the hysteresis window allow to achieve higher energy savings at the price of a higher frequency of switching on/off APs. Such hysteresis window must therefore be dimensioned trading off these two effects and to this purpose the $M/M/\infty$ model can be exploited. Whereas the high threshold should be statically fixed according to Quality of Service (QoS) constraints, in order to always guarantee a minimum amount of available bandwidth per user, the low threshold (and therefore the hysteresis window width) can be dynamically set in order to find a proper trade-off between the energy savings and the frequency of the switching on/off operation. The $M/M/\infty$ queuing system can be seen as a continuous time Markov chain, $X(t) = \{i = 0, 1, \cdots\}$, in which state i represents the number of users in the system; arrivals at rate λ correspond to users accessing the WiFi, departures represent users finishing sessions. In the Markov chain, the transition from any state i to state $i+1$ occurs at rate λ, while the transition from i to state $i-1$ occurs at rate $i\mu$. While the $M/M/\infty$ queue can represent the users access to the network, it is not adequate to derive some performance indicators that are associated to the hysteresis window mechanisms, such as the number of active APs. In particular, when i users are in the system, with $T_h^{(k)} - \omega < i < T_h^{(k)}$, we can either have k or $k+1$ active APs, depending on the past behavior, i.e., depending on whether the $(k+1)^{th}$ AP was switched on without having been switched off, or vice-versa. When the evaluation of the number of active APs is necessary, a Markov chain definition based on the number of users only is not enough; the Markov chain state definition should be enhanced with a state variable representing the number of active APs. We, thus, define a new Markov chain with state (i,k), where k represents the number of active APs. A portion of the Markov chain is represented in Fig. 12 and the transitions from state (i,k) are reported in Tab. III. For the solution of the MC we use standard techniques after truncating the infinite state MC to values whose probability is below 10^{-9}.

C. AP operation

We now discuss how the model can be used to understand AP operation under the RoD strategies. Data obtained from our traces show that the value of the arrival rate, although changing over days, is pretty constant over a whole single day, during the steady-state phase; λ ranges between a minimum value of 0.0565 [user/second] to a maximum value of 0.1116 [user/second]. This means that our $M/M/\infty$ model can be applied using a different value of λ for each day, and that the mean number of users in the system (ρ=λ/μ) ranges between 52.7 and 101.8 users. The study of the continuous time Markov chain modeling allows us to investigate the behavior of the system when the described RoD strategy is applied under different combinations of $T_h^{(k)}$ and ω. First, we focus on the effect of these parameters on the patterns of AP activity. The following indicators are evaluated:

- **Active time**: average duration of a period of activity of the AP. Given that there are k active APs, it can be defined as the time elapsed from the instant in which a new AP, the $(k+1)^{th}$ AP, is switched on to the instant in which that AP is switched off again. In the Markov chain, such time can be computed as the *average first passage time* [26] from state $(T_h^{(k)}, k+1)$ to state $(T_l^{(k+1)}, k)$.
- **Off time**: average time spent in off state by an AP, that is also the time needed for an AP that has been switched off to be switched on again. In this case, it can be computed as the *average first passage time* from the state $(T_l^{(k)}, k-1)$ to state $(T_h^{(k-1)}, k)$ in the Markov chain.

The values of active and off times are important to properly set the hysteresis window width. Indeed, the energy saving

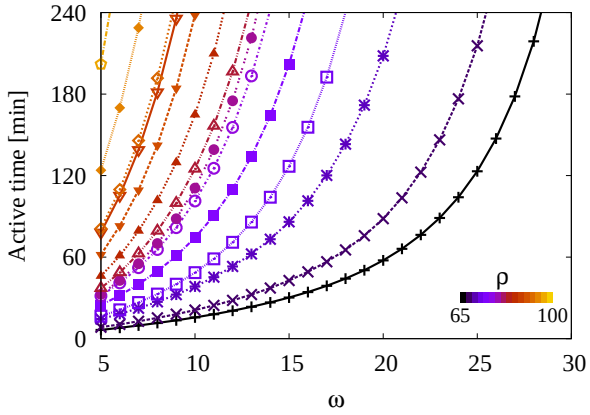


Fig. 13: Average active time versus hysteresis window width for different values of ρ, fixed $T_h^{(k)}$ = 80 users; teaching period.

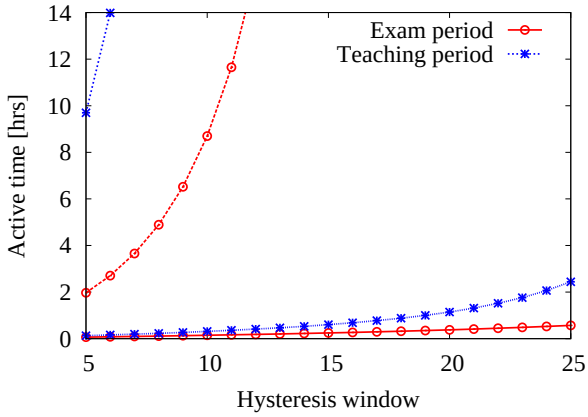


Fig. 14: Average active time (for the maximum and the minimum values of ρ) versus the hysteresis window width with fixed $T_h^{(k)}$ = 80 users during the teaching period (blue curves) and the exam period (red curves).

is proportional to the duration of the off time and off and active times affect the switching frequency of the AP, and, hence, the stability of the system. In what follows, we show and discuss the active and off times only for the AP that is the most frequently switched on and off. Indeed, when the system is operating with N APs at a given value of load $\rho = \lambda/\mu$, the number of active APs varies in a range that is quite limited; some APs are always on and one AP switches on and off to absorbs peaks of traffic. For example, if the mean number of users is 80, and M is set to 20, so that the high thresholds are set to $M \cdot k = 20 \cdot k$, three APs are always active and a fourth AP is dynamically switched on and off following demand fluctuations. The active time and the off time have been evaluated under different combinations of $T_h^{(k)}$ (ranging from 50 to 100 users) and ω (from 5 to 30 users) and under different mean values of the number of users in the system, corresponding to different days in the teaching and exam periods. In general, the days in the teaching period show higher values of ρ, in the range [68.6, 101.8] users, with respect to the exam period, in the range [52.7, 94.8] users. Given ρ, the values of $T_h^{(k)}$ and ω influence the duration of the active time. We start by focusing on the effect of ω. Figs. 13 and 14 show the active time for increasing values of ω, when $T_h^{(k)}$ = 80 users. Different curves refer to different days, i.e., different values of ρ, of either the exam or the teaching period. In Fig. 13 (teaching period) the highest and lowest curves correspond to those days whose values of ρ are the minimum and the maximum one, respectively. The AP active time grows with ρ, because the higher capacity demand calls for longer activity periods of the AP. The active time grows also with ω, since to reduce the switching frequency the hysteresis window makes AP activity times longer. Confirming this, the average active time tends to increase even more slowly with ω in the case with $\rho \approx 57$ users, corresponding to the day with a low value of ρ, that is registered during the exam period. This can be observed as well in Fig. 14, where the curves of the average active time for the maximum and the minimum values of ρ are plotted both for the teaching period (blue curves) and for the exam period (red curves).

The duration of the active time should be larger than the minimum time needed for an AP to complete the switching on process. In particular, for a given value of ρ and a fixed $T_h^{(k)}$, settings of ω determining an active time shorter than 5 minutes (a conservative estimation of the AP switching time) should be avoided.

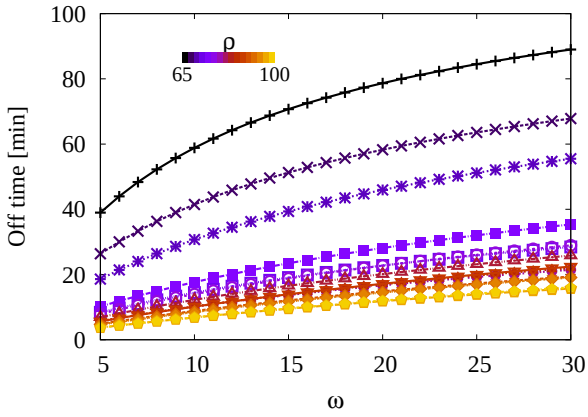


Fig. 15: Average off time versus hysteresis window width for different values of ρ, fixed $T_h^{(k)}$=80 users; teaching period.

In a similar way to the active time, the off time is affected by the values of $T_h^{(k)}$, ω and ρ. Fig. 15 shows the off time versus the hysteresis window width for different values of ρ given a fixed $T_h^{(k)}$ = 80 users. Also in this case, the off time increases as ω grows larger. However, the growth is much slower than for the active time, with the off time being smaller than one hour in most of the cases. Moreover, the off time increases as ρ decreases. The evaluation of the off time is particularly important since the energy savings are proportional to its duration. Furthermore, also in this case, the off time should be long enough (e.g., a conservative 5 minutes) to allow the AP to fully switch off; the minimum value of ω for a given setting of $T_h^{(k)}$ and ρ should be defined accordingly.

We now consider the average number of switching events over the steady-state period, that we denote by e_{ss}, and that

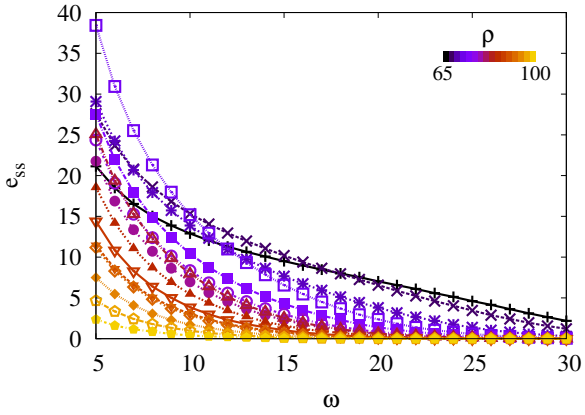


Fig. 16: Average number of switching events over the steady-state period with respect to the hysteresis window width for different ρ with fixed $T_h^{(k)}$ = 80 users; teaching period.

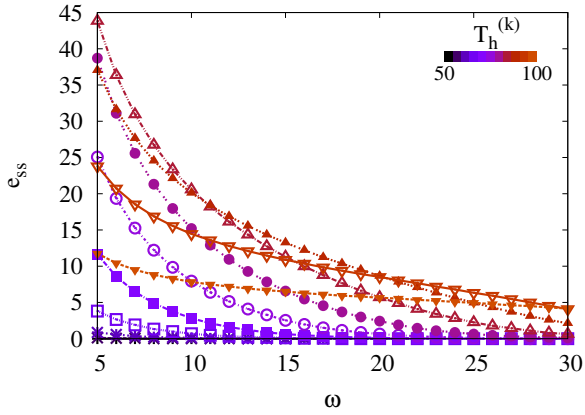


Fig. 17: Average number of switching events over the steady-state period with respect to the hysteresis window width for different values of $T_h^{(k)}$ with fixed ρ = 85 users.

can be defined as the number of switching on and switching off operations performed over the steady-state period (8 hours from 10:00 to 18:00). Again, we will consider only the switching events of the AP devoted to absorb peaks of demand. The frequency of switching on events per second (f_{on}) can be computed as:

$$f_{on} = \Pi_{(T_h^{(k)}-1,k)} \cdot \lambda \tag{2}$$

where $\Pi_{(T_h^{(k)}-1,k)}$ is the steady-state probability of being in the state $(T_h^{(k)}-1,k)$ and λ is the arrival rate. The frequency of switching off operations per second (f_{off}) is given by:

$$f_{off} = \Pi_{(T_l^{(k+1)}+1,k+1)} \cdot (T_l^{(k+1)}+1) \cdot \mu \tag{3}$$

where $\Pi_{(T_l^{(k+1)}+1,k+1)}$ is the steady-state probability of being in the state to $(T_l^{(k+1)}+1,k+1)$ with the AP switched on and μ is the service rate. Although there are additional switching events during the ramp-up and the emptying period, only the steady-state period has been considered, since our model is describing the system in the steady-state phase. Combining (2) and (3), the average number of switching events over the steady-state period, e_{ss}, for the considered AP is:

$$e_{ss} = \left(f_{off} + f_{on}\right) \cdot (60 \cdot 60 \cdot 8) \tag{4}$$

where the last factor translates a frequency in seconds into number of events in the steady-state period. Fig. 16 reports the values of e_{ss}, (4) versus increasing values of ω for $T_h^{(k)}$ = 80 users and for different values of ρ. In general, as ω becomes larger, the e_{ss} decreases. However, for values of ρ in a range higher than the considered $T_h^{(k)}$ (for ρ > 80 users), as ρ increases the e_{ss} decreases. On the contrary, for values of ρ lower than 80 users, for small values of ω, e_{ss} decreases with the hysteresis window width more slowly than it does for wider values of ω; as a consequence, in this case e_{ss} increases with ρ only for large values of ω, whereas for small ω values, e_{ss} decreases as ρ increases. Fig. 17 shows the variation of e_{ss} versus ω and for increasing values of $T_h^{(k)}$, considering a fixed ρ of 85 users. For very high or very low values of $T_h^{(k)}$ with respect to the values of ρ, e_{ss} is close to 0 events over the steady-state period for any value of ω.

VI. Performance evaluation

This section introduces some results about energy saving that can be obtained with the application of the RoD strategies. Some scenarios are considered to perform a comprehensive analysis of the system parameter setting, to trade-off energy saving and system stability and, finally, to assess the impact of RoD strategies on QoS provided to users.

A. Parameter setting for energy saving

We now focus on the impact of the RoD parameters on energy saving. With respect to the case in which all the available APs are always kept on regardless the users' demand, variable amounts of energy saving can be achieved when the RoD strategy is applied to a given dense WLAN scenario. Without any RoD strategy all APs are always active and consuming energy, while under a RoD strategy this is what may happen in the steady-state period in terms of energy consumption:

- there may be some APs that are constantly on providing capacity for normal usage and they are responsible of a *constant amount of energy consumption*;
- some APs are constantly in an off state, since they are not needed according to the current system load; these APs do not consume energy, providing a *static energy saving*;
- one or more APs providing capacity for peaks of service demand are switched on or off from time to time following the users' demand and thus allowing a *dynamic energy saving*.

The analysis of the results about the active/off time and the average number of switching events over the steady-state period is important for quantifying energy saving. In particular, for a given scenario, the energy that can be saved thanks to the RoD strategy highly depends on the setting of $T_h^{(k)}$ and ω but also on the variability of the load of the system in terms of average number of users (the users' demand).

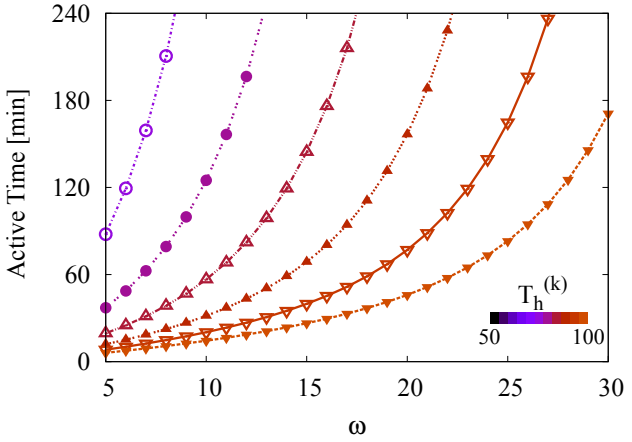


Fig. 18: Active time with respect to ω for different $T_h^{(k)}$.

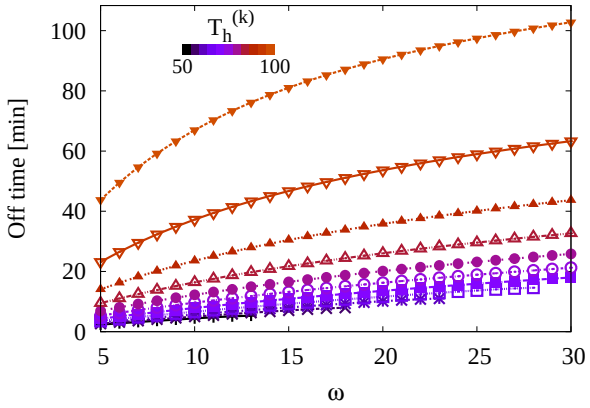


Fig. 19: Off time with respect to ω for values of $T_h^{(k)}$.

Fig. 18 shows the active time versus ω for different values of $T_h^{(k)}$, considering a single day (i.e., for a constant value of $\rho \approx 87$ users). As ω grows, the active time becomes larger; also, for decreasing values of $T_h^{(k)}$, higher values of the active time are observed. Furthermore, for $T_h^{(k)}$ much higher than the considered ρ, the curves representing the active time tend to be lower and to grow more slowly; on the contrary, for $T_h^{(k)}$ much smaller than the considered value of ρ, the curves representing the active time are higher and show a steeper growth. The off time exhibits an opposite behavior, as it can be seen from Fig. 19 showing the off time versus ω for different values of $T_h^{(k)}$, and constant values of $\rho \approx 87$ users. The off time decreases with decreasing values of $T_h^{(k)}$ and slowly increases while higher values of ω are set. Finally, for $T_h^{(k)}$ much higher than the considered ρ, the curves representing the off time grow slightly faster, whereas for $T_h^{(k)}$ smaller than the considered ρ, the active time shows a very flat growth.

These results indicate that the finest regulation of energy saving can be obtained on those APs operating with a value of $T_h^{(k)}$ which is pretty close to the average number of users in the system. Therefore, if the value of $T_h^{(k)}$ based on which an AP is switched on is much higher than the current average number of users in the system, the active time is small and the off time is large, meaning that the AP remains off most of the time without significantly contributing to the energy consumption of the system. In an opposite way, if the value of $T_h^{(k)}$ based on which an AP is switched on is much lower than the current average number of users in the system, the active time is long and the off time short, meaning that that AP will be in an on state most of the time and very unlikely it will be turned off, thus contributing to the constant portion of energy consumption during the considered steady-state period (*constant energy consumption*) and no energy saving can be obtained in this case. When the $T_h^{(k)}$ value is closer to the average number of users in the system the active time and the off time will show intermediate values allowing that AP to be switched on and off from time to time and the tuning of ω may enable to regulate the percentage of energy saving that can be obtained in relation to that AP with respect to the configuration with all the AP always on (*dynamic energy saving*). This is also confirmed by the values of the the average number of switching events over the steady-state period plotted with respect to all the possible $T_h^{(k)}$ values with fixed ω and for different values of ρ shown in Fig. 20. The e_{ss} shows a peak value for each ρ right in correspondence of those $T_h^{(k)}$ values close to ρ: the AP in that case is switched on or off frequently and during the off state energy can be dynamically saved. For $T_h^{(k)}$ higher or lower than ρ the the average number of switching events over the steady-state period tends to progressively decrease: in the first case because the AP is rarely turned on (no considerable amounts of energy are consumed), whereas in the latter case because the AP remains on most of the time (no significant amounts of energy can be saved in this case).

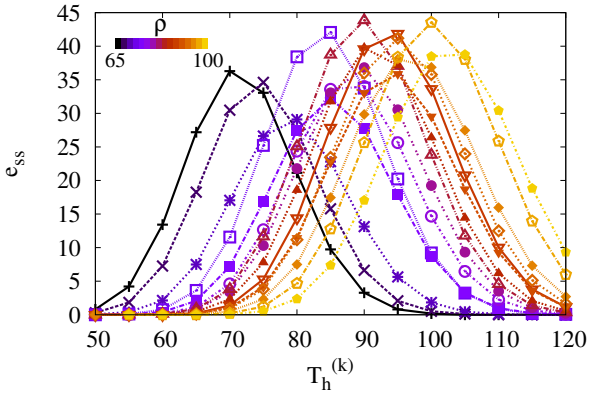
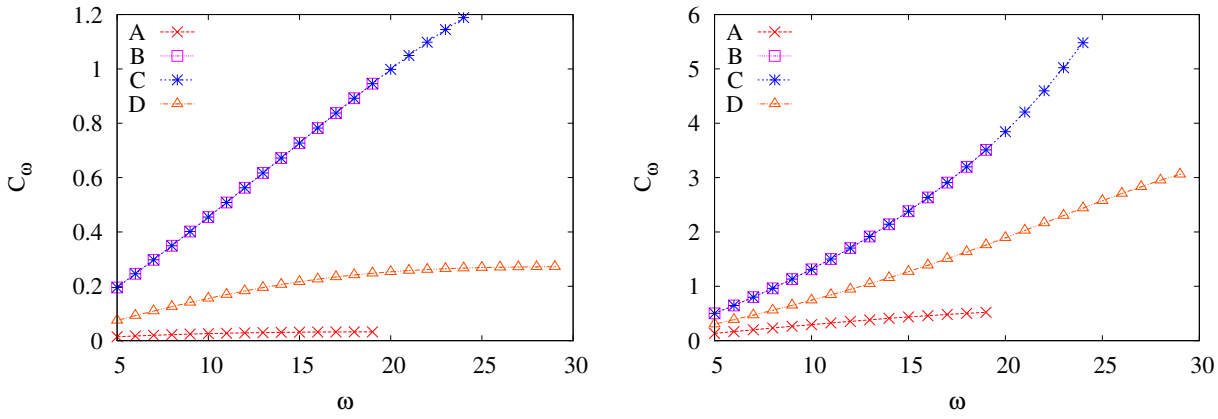


Fig. 20: Average number of switching events over the steady-state period with respect to increasing values of $T_h^{(k)}$ for different ρ with fixed $\omega = 5$ in the teaching period.

B. Trading off energy consumption and system stability

We investigate now the trade-off between energy saving and system stability. To do so, some specific scenarios are considered. We consider different values of the number of users handled by each AP, typically set by the system administrator depending on the desired level of QoS. Different values of the



(a) $\rho \approx 97$ users (b) $\rho \approx 83$ users

Fig. 21: Energy cost due to the hysteresis window in the four scenarios versus ω in two different days.

number of available APs have also been considered, resulting in the following four scenarios:

- scenario A: M=20 users, N=5 APs;
- scenario B: M=20 users, N=6 APs;
- scenario C: M=25 users, N=5 APs;
- scenario D: M=30 users, N=4 APs.

These settings translate into different values of $T_h^{(k)} = kM$ for each scenario. To investigate the trade-off between system stability and energy saving under the RoD strategy, we quantify the additional energy that is consumed by introducing the hysteresis window with respect to the case in which there is no hysteresis window (ω=0), that represents the minimum cost case. We focus on the AP that provides peak capacity and is responsible of the dynamic energy saving; let us denote it by k. We define the energy cost due to the hysteresis window, C_ω, as the ratio between the fraction of time during which the AP is on and it is in the hysteresis window, F_ω, over the fraction of time during which the AP is on when ω=0, F_{on}:

$$C_\omega = \frac{F_\omega}{F_{on}} \tag{5}$$

From the Markov chain we can compute:

$$F_{on} = \sum_{i=T_h^{(k-1)}}^{\infty} \Pi_{(i,k)} \tag{6}$$

and

$$F_\omega = \sum_{i=T_h^{(k-1)}-\omega+1}^{T_h^{(k-1)}-1} \Pi_{(i,k)} \tag{7}$$

F_{on} in (6) represents a cost for the AP being on, independently on the hysteresis window, while in (7) F_ω is the cost during the periods in which the system is *in* the hysteresis window.

First, a day with a high value of ρ ($\rho \approx 97$ users) has been analyzed. Being the load quite high, $N-1$ APs are active most of the time for all the scenarios. Therefore, the N^{th} AP is the one providing peak capacity and being responsible of the dynamic energy saving. The values of the energy cost C_ω, (5), in the four scenarios are reported in Fig. 21a. Clearly, C_ω increases with ω and it ranges from a minimum of 0.014 (only 1.4% of additional cost) to a maximum of 1.189 (more than 100% additional cost). The highest values of C_ω are observed in the scenarios B and C, both with $T_h^{(N-1)}$=100 users, that is very close to the mean number of users. Finally, a faster growth of C_ω with ω is observed for increasing values of $T_h^{(N-1)}$. These results have been compared with those obtained for the same AP in the same scenarios for a lower load, namely, $\rho \approx 83$ users, see Fig. 21b. The behavior is similar to the previous case with a cost to be paid for introducing the hysteresis window that is larger in relative terms (up to 5 times). Notice, however, that in absolute terms the probability that the AP is active is much lower due to the low user demand. The width of ω should therefore be set carefully, since it affects the cost in terms of additional energy consumption of the system. The same scenarios have been considered to estimate the average daily energy consumption of the whole system under the RoD strategy and to compare it with the average daily energy consumption without any RoD algorithm. In particular, the average energy saving per week day has been computed in percentage as:

$$S = \left(1 - \frac{E_{RoD}}{E}\right) \cdot 100 \tag{8}$$

where E_{RoD} is the energy consumption of all APs in the cluster under the RoD strategy whereas E is the energy consumed when all APs are always kept active. Table IV reports the average energy saving per working day defined in (8) in two cases ($\rho \approx 83$ users and $\rho \approx 57$ users) under the four scenarios. The average energy consumption is significantly reduced in all cases with respect to the absence of any RoD strategy (up to 59.1% of power can be saved) and, clearly, the reduction is more evident for small values of ω. Furthermore, the percentage of saving that can be achieved is larger when the average number of users is lower; in case of $\rho \approx 57$ users (corresponding to the exam period), saving is up to 59.1% in scenario B. These results are very promising in perspective, in view of the possible deployment of the RoD strategy on the whole Campus, especially considering that in this study we were focusing on periods of high activity in the Campus.

TABLE IV: Average energy saving (%) per working day

Scenarios		Hysteresis window (ω) 5	10	15	20	25
ρ = 83	A	36.74	35.6	34.55		
	B	46.8	45.6	44.35		
	C	44.4	44.1	43.7	43.1	
	D	41.7	40.3	38.6	36.61	34.4
ρ = 57	A	51.2	49.5	47.7		
	B	59.1	57.5	55.83		
	C	57.5	56.64	56	55.6	
	D	53.5	51.5	49.2	47.2	46.13

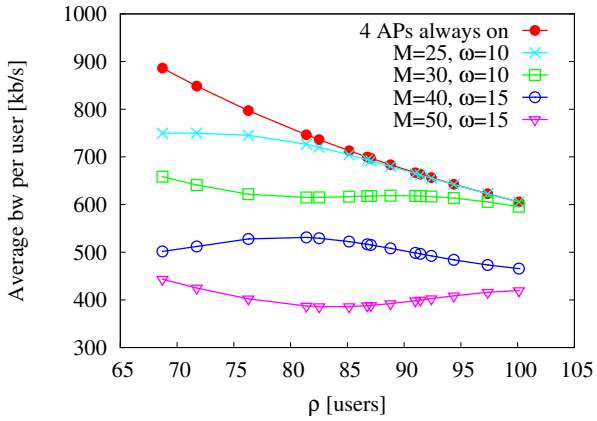


Fig. 22: Average bandwdith per user, in $N = 4$ APs scenarios.

During summer, exam periods, vacations, the activity in the Campus is reduced and the expected savings are even larger.

C. Available bandwidth per user

In the previous sections, we mainly focused on energy saving and system stability. We now consider the potential impact of RoD strategies on the QoS provided to users that we evaluate in terms of average available bandwidth per user, denoted as B_u, that can be computed as:

$$B_u = \sum_{i=0}^{\infty} \sum_{k=1}^{N} \frac{\Pi_{(i,k)} \cdot k \cdot B_{AP}}{i} \qquad (9)$$

where B_{AP} is the maximum throughput per AP. Clearly, due to protocol overheads and access protocol characteristics, the maximum throughput that can be achieved with an AP is much lower than the maximum nominal bit-rate, which in our case is equal to 54 Mb/s. Thus, to estimate the value of B_{AP} we performed several tests on the APs usage in our Campus and estimated a value of B_{AP} equal to 15 Mb/s.

Fig. 22 shows the results obtained in a scenario in which up to $N = 4$ APs can be activated. Different values of the parameters M and ω are considered. The red curve represents the available bandwidth for the always on case, i.e., when all the 4 APs are always active. The figure shows that when RoD strategies are adopted, the available bandwidth per user is smaller than in the all-on case. However, for low values of M, the average bandwidth per user is not considerably smaller: the RoD strategy is not aggressive, APs are switched on when relative few users are in the system. Conversely, the cases of large values of M correspond to cases in which it

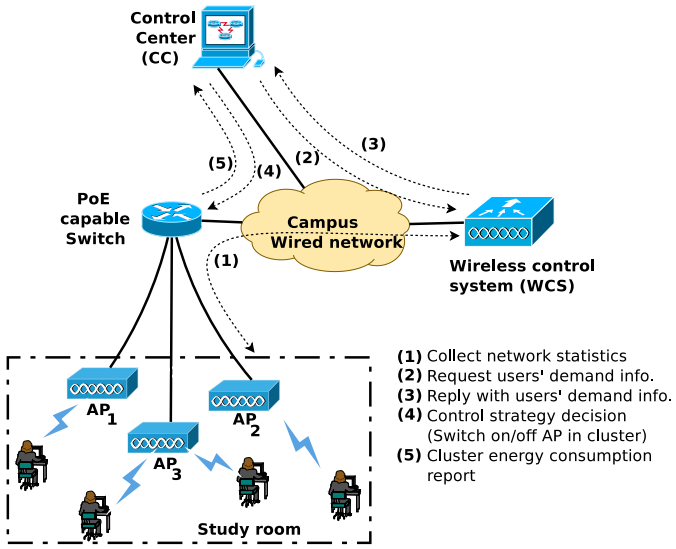


Fig. 23: Trace collection and experimental environment.

is accepted that many users share the same AP, and, hence, the available bandwidth per user results smaller than in the all-on case. Notice, however, that the setting of M is defined by the network administrator that chooses, as QoS constraint, the minimum value of available bandwidth that is acceptable. The figure shows that once this constraint is defined through the setting of M, the setting of ω and the value of the load ρ do not have a significant impact on it and the QoS constraint can be easily met.

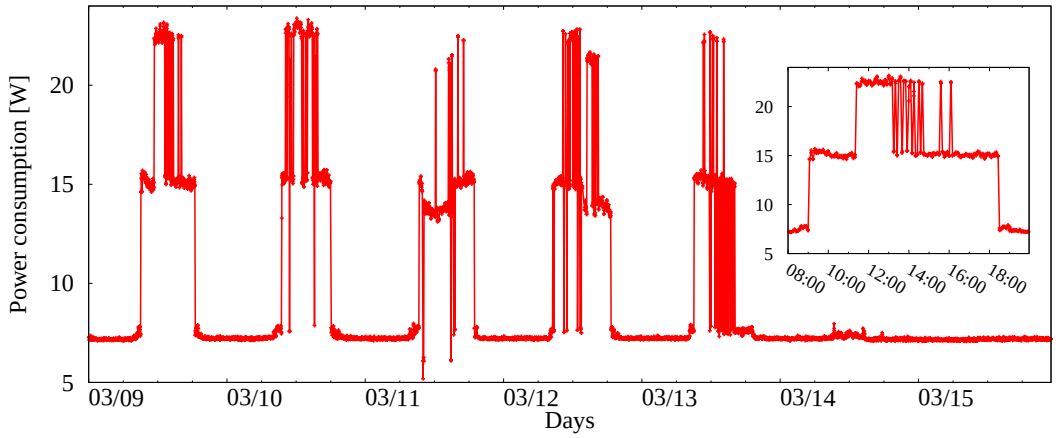
VII. Experimental Setup

As detailed in Section V, one of the outcomes of RoD design is the threshold settings. These settings determine the energy saving efficiency, the APs switching frequency, the stability of the network and QoS provided to users. As a proof of concept, in this section we demonstrate the effectiveness and impact on QoS of RoD strategies using a production network.

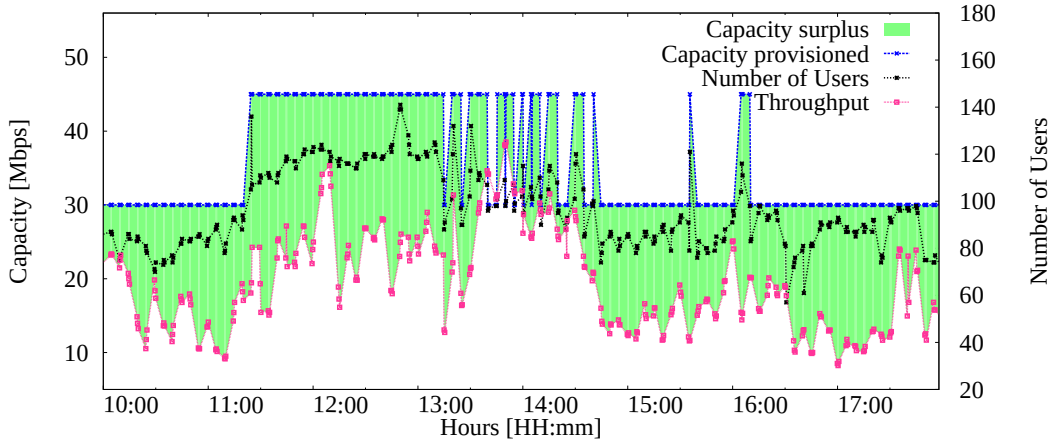
A. Experimental results

We perform experiments in the study room from which we collected our traces. The experimental set up is shown in Fig. 23. The room is equipped with 3 APs which are connected to a PoE capable switch that enables us to remotely control the switching on/off of the APs; based on threshold settings, from a control center (CC) which runs the RoD algorithm. Once a decision is made by the RoD algorithm, the CC controls (switches on/off) the PoE ports that power the APs by sending an snmpset command to the switch. The CC makes its decision based on session information obtained from the Cisco Wireless Control System (WCS) [27]; a management system that receives associated clients session statistics and that controls the PoliTo WiFi network. The CC polls the WCS for this information every 5 minutes, see Section VII-B for a discussion on the impact of the polling time.

Fig. 24a depicts the power consumption of the testbed for the following RoD settings: $M = 50$ and $\omega = 10$. The figure



(a) Daily power consumption of a production network.



(b) Capacity provisioned and throughput (March 16, 2015).

Fig. 24: Testbed power consumption and available network capacity.

shows that our threshold based RoD strategy well adapts the network capacity to clients demand. The small inset figure zooms in a specific day (March 9th) power consumption. It can be seen that in the idle period, during weekends and overnight, only one AP is required to satisfy the demand since there are few or no users in the system. During ramp-up and emptying periods, early in the morning and late in the afternoon, as the demand increases/decreases the RoD algorithm switches on/off the second AP to adapt the capacity to the demand. In the steady state (10:00 - 18:00), the capacity adaptation is more dynamic since the number of users in the study room varies around the typical value. The average energy saving over a week period, computed as in (8), is about 58.6%. This saving increases to 68.9% during weekends. The results confirm the saving that was already estimated in the previous section and the potential of the RoD approach when applied and extended to the whole Campus.

As discussed in the previous section, by reducing the available capacity, RoD strategies have an impact on the available bandwidth per user and, hence, a potential effect on the QoS provided to users. While it is not possible to estimate the demand of capacity from the users, to assess this impact, Fig. 24b compares the available network capacity with the throughput for the RoD strategy deployed in our testbed; the number of associated users is also reported. The throughput is estimated as the ratio of the aggregate traffic exchanged by the APs over a 5 minute time windows. The figure clearly shows that the network capacity follows the number of associated users and the provisioned capacity is always quite larger than the throughput, meaning that probably the users do not undergo quality of service degradation during the experiment. In general, the QoS perceived by the users depends also on the characteristics of the generated traffic, its volume and its burstiness.

To investigate these characteristics, we have analyzed the amount of traffic transmitted and received by the APs with a single user. These results are shown in Fig. 25, in which we have computed the CDF of the data per session. The figure refers to the five working days of a single week, each of them is represented by a line of a different color. Many sessions carry very little traffic: 40% of the sessions exchange less than 100 KB of data. Moreover, most of the data exchanged during a single session (i.e., more that 95%) has a size that is smaller than 100 MB. It is important to remark that this is the volume of data sent and received by a single user with the WiFi network. Combining these values with the quite long

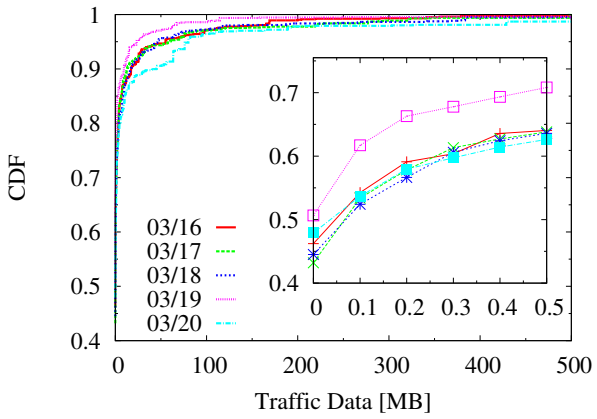


Fig. 25: CDF of the data traffic per session.

session durations that were observed in Sec. IV, it emerges that, in our Campus scenario, the capacity demand per user is relatively low, since users generate little traffic and many users that are associated in the same moment do not generate traffic. In other scenarios, the situation might be different. Hence, the evaluation of QoS and the proper setting of the parameters needed to meet QoS constraints should be specifically defined scenario by scenario.

B. Implementation issues

We now discuss possible implementation issues, drawbacks and potential critical aspects of the deployment of RoD strategies.

1) Client disconnection: The principle behind a RoD strategy is to scale network resources to clients' capacity demand. In a dense WLAN where multiple APs provide similar coverage, RoD strategies switch on/off APs to match the time varying clients' capacity demand with the network capacity. When switching off an AP, there is the possibility that sessions connected through the AP are abruptly interrupted, causing some service deterioration. Thus, before switching off an AP, clients should be gracefully roamed to other active APs. This can be done in different ways depending on the actual implementation of the system. One possibility, whenever the implementation allows it, is that the network can proactively request a user to handover to another AP. Alternatively, the AP that is supposed to be switched off should undergo a gradual AP transmission power reduction that encourages clients to move to neighboring APs with better SNR. Once clients have roamed to neighboring APs, an AP can be switched off thus avoiding client disconnection. In our testbed scenario, since the WCS implements proprietary solutions, we could not change the configuration and operation of the WCS and, for this reason, we could not implement none of the solutions mentioned above.

2) System polling interval: The WCS monitors and updates the state of the network in real time. As soon as a user associates or disassociates, the WCS updates the network status information. Switching decisions can thus be taken instantaneously, as soon as the number of associated users goes above or below the predefined thresholds, only if the switching decisions are taken by the WCS. As previously mentioned, in our experimental setup we could not integrate the switching decision engine into the WCS and we had, instead, to make use of a CC that polls the WCS status according to some polling intervals that we defined to be equal to 5 minutes. The rational behind the choice of a 5 minutes polling interval is the following. In our implementation, the AP requires between 2 to 3 minutes to transit from a complete shutdown state to a full active state. Furthermore, to start the service, the APs need to register to the WCS which takes one or two more minutes. Thus, the switching on frequency should not be higher than a possible switch on every 4 to 5 minutes. Given this physical limitation, the CC uses the default 5 minutes SNMP polling window to collect the required statistics from the WCS.

3) AP lifetime: One of the common concerns about RoD strategies based on device switching is that frequent status changes might be harmful to the device, possibly causing some reduction of its lifetime. There exists in the literature some models for estimating the number of cycles to fail of an electronic device; for example, the very well-known Coffin-Manson model describes how repeated thermal cycles induce a cyclical stress that tends to weaken materials and has been successfully used for mechanical failure, material fatigue failure or material deformation [28]–[31]. To the best of our knowledge no specific study has been done for WiFi APs. However, in our experience with the experimental activity, that has been going on for more than 1.5 years and that requires an average number of switching on/off events per AP per day of about 3, we never experienced device tear down or any technical glitches as a result of AP switching on/off. Notice also that the results presented in the paper show that, through proper parameter setting, the number of switching events can be controlled and maintained small.

VIII. Conclusions

Resource on-demand (RoD) is a viable approach to save energy in dense WLANs where large number of APs are deployed to have enough capacity to satisfy peak users' service demand. RoD strategies scale the network capacity to the demand by switching off extra capacity during low load period and switching on APs as the demand increases.

The effectiveness of a RoD strategy depends on its parameters setting; and, in particular, on the AP activation/deactivation thresholds. These settings determine the stability of the network, the amount of achievable saving and the QoS provided to users. To analyze the right settings of these parameters and its impact on network performance, we modeled a dense campus WLAN with the model parameters derived from production network traces. We also verified the model by fitting its behavior to the actual behavior of the production network. Based on the inferred model, we presented a detailed investigation on RoD threshold and hysteresis window settings and we derived guidelines of their setting. We also conducted experiments on the production network from which we collected traces as a proof of concept of RoD strategy and discussed some implementation issues. The analysis has been

done for two different academic periods; teaching and exam periods, to help us provide an all inclusive observation of the system.

From the analysis, through both the model and the experiments, we observed that high values of the AP activation threshold imply large energy saving, while the hysteresis window width should be carefully set so as to jointly guarantee high saving, a smooth network operation and desired levels of available bandwidth per user.

The estimated and tested saving is large and the practical implementation of the RoD approach is simple and does not require any specific hardware facilities. The approach is therefore extremely promising. This study opens the pave to extend the approach to the whole Campus. On a large scale, the savings are expected to be even more significant, since there are large areas of the Campus that are less used than the considered study room, and there are long periods of vacations with very low network utilization.

REFERENCES

[1] W. Lemstra, V. Hayes, and J. Groenewegen, *The innovation journey of Wi-Fi: The road to global success*. Cambridge University Press, 2010.

[2] P. Leisching and M. Pickavet, "Energy footprint of ICT: Forecasts and network solutions," in *OFC/NFOEC 09, Workshop on Energy Footprint of ICT: Forecast and Network Solutions*, March 2009.

[3] A. P. Jardosh, K. Papagiannaki, E. M. Belding, K. C. Almeroth, G. Iannaccone, and B. Vinnakota, "Green WLANs: On-demand WLAN Infrastructures," *Mobile Networks and Applications*, December 2009.

[4] Cisco Systems, "Cisco Aironet 1200 Series Datasheet," January 2010. [Online]. Available: http://www.cisco.com/c/en/us/products/collateral/wireless/aironet-1200-access-point/produc t_data_sheet09186a00800937a6.html

[5] F. Debele, N. Li, M. Meo, M. Ricca, and Y. Zhang, "Experimenting Resource-on-Demand Strategies for Green WLANs," in *Greenmetrics workshop*. ACM, 2014.

[6] C. Phillips and S. Singh, "Analysis of WLAN Traffic in the Wild," in *NETWORKING 2007. Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet Lecture Notes in Computer Science*. Springer, 2007, pp. 1173–1178.

[7] S. Geirhofer, L. Tong, and B. Sadler, "Dynamic spectrum access in WLAN channels: empirical model and its stochastic analysis," in *TAPAS '06 Proceedings of the first international workshop on Technology and policy for accessing spectrum*. ACM, 2006.

[8] C. Na, J. Chen, and T. Rappaport, "Measured Traffic Statistics and Throughput of IEEE 802.11b Public WLAN Hotspots with Three Different Applications," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, 2006.

[9] F. Hernandez-Campos, M. Karaliopoulos, M. Papadopouli, and H. Shen, "Spatio-temporal modeling of traffic workload in a campus WLAN," in *WICON '06 Proceedings of the 2nd annual international workshop on Wireless internet*. ACM, 2006.

[10] M. Karaliopoulos, M. Papadopouli, E. Raftopoulos, and H. Shen, "On scalable measurement-driven modeling of traffic demand in large WLANs," in *15th IEEE Workshop on Local & Metropolitan Area Networks, 2007, LANMAN 2007*, 2007, pp. 102–110.

[11] E. Halepovic, M. Ghaderi, and C. Williamson, "On the Performance of Redundant Traffic Elimination in WLANs," in *ICC 2012*. IEEE, 2012.

[12] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *IEEE International Conference on Communications Workshops*, 2009, pp. 1–5.

[13] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 74–79, 2010.

[14] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, 2010.

[15] K. Abdallah, I. Cerutti, and P. Castoldi, "Energy-efficient coordinated sleep of lte cells," in *IEEE International Conference on Communications (ICC)*, 2012, pp. 5238–5242.

[16] K. Adachi, J. Joung, S. Sun, and P. H. Tan, "Adaptive coordinated napping (conap) for energy saving in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5656–5667, 2013.

[17] L. Budzisz, F. Ganji, G. Rizzo, M. Ajmone Marsan, M. Meo, Y. Zhang, G. Koutitas, L. Tassiulas, S. Lambert, B. Lannoo *et al.*, "Dynamic resource provisioning for energy efficiency in wireless access networks: a survey and an outlook," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2259–2285, 2014.

[18] R. Ding and G.-M. Muntean, "Device characteristics-based differentiated energy-efficient adaptive solution for video delivery over heterogeneous wireless networks," in *IEEE WCNC*, 2013.

[19] A. da Silva, M. Meo, and M. A. Marsan, "Energy-performance trade-off in dense WLANs: A queuing study," *Computer Networks*, July 2012.

[20] J. Lorincz, A. Capone, and M. Bogarelli, "Energy savings in wireless access networks through optimized network management," in *IEEE ISWPC*, May 2010, pp. 449–454.

[21] N. Mishra, K. Chebrolu, B. Raman, and A. Pathak, "Wake-on-WLAN," in *WWW*. IW3C2, May 2006.

[22] Y. Kondo, H. Yomo, S. Tang, M. Iwai, T. Tanaka, H. Tsutsui, and S. Obana, "Energy-efficient WLAN with on-demand AP wake-up using IEEE 802.11 frame length modulation," *Computer communications*, vol. 35, no. 14, 2012.

[23] M. Ajmone Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "A simple analytical model for the energy-efficient activation of access points in dense WLANs," in *e-Energy*. ACM, 2010, pp. 159–168.

[24] M. S. Gast, *802.11ac: A Survival Guide*. OReilly, 2013.

[25] Sendra, Sandra and Garcia, Miguel and Turro, Carlos and Lloret, Jaime, "WLAN IEEE 802.11 a/b/g/n Indoor Coverage and Interference Performance Study," *International Journal on Advances in Networks and Services*, vol. 4, no. 1 and 2, pp. 209–222, 2011.

[26] L. Lipsky, *Queueing Theory: A Linear Algebraic Approach*, 2nd ed. Springer Publishing Company, Incorporated, 2008.

[27] Cisco Systems, "Cisco wireless control system configuration guide," June 2010. [Online]. Available: http://www.cisco.com/

[28] G. Yang, *Life Cycle Reliability Engineering*. Wiley, 2007.

[29] H. Cui, "Accelerated temperature cycle test and coffin-manson model for electronic packaging," in *Reliability and Maintainability Symposium, 2005. Proceedings. Annual*, 2005, pp. 556–560.

[30] R. C. Blish *et al.*, "Temperature cycling and thermal shock failure rate modeling," in *Reliability Physics Symposium, 1997. 35th Annual Proceedings., IEEE International*. IEEE, 1997, pp. 110–117.

[31] N. Durrant and R. Blish, "Semiconductor device reliability failure models," *http:/www. sematech. org*, 2000.

Fikru Getachew Debele received Master of Technology from Indian Institute of Technology (IITK) in Electrical Engineering in 2005. From 2006 to 2008 he was working in ArbaMinch University, Ethiopia as an ICT coordinator until he joined Politecnico di Torino in January 2009 to pursue his postgraduate study. He completed his Ph.D. degree in Electronics and Communication Engineering in 2012. Since January 2013, he has been working with the Telecommunication Networks group as a PostDoc researcher at Politecnico di Torino. His area of interests are in the field of: energy efficient networks, software defined networking (SDN) and smart grids.

Michela Meo received the Laurea degree in Electronic Engineering in 1993, and the Ph.D. degree in Electronic and Telecommunications Engineering in 1997, both from the Politecnico di Torino, Italy. Since November 2006, she is associate professor at the Politecnico di Torino. She co-authored about 200 papers and edited a book with Wiley and six special issues of international journals, including ACM Monet, Performance Evaluation, and Computer Networks. She chairs the Steering Committee of IEEE OnlineGreenComm and the International Advisory Council of ITC. She was program co-chair of several conferences among which ACM MSWiM, IEEE Online GreenComm, IEEE ISCC, IEEE Infocom Miniconference, ITC. Her research interests include the field of performance evaluation and modeling, green networking and traffic classification and characterization.

Daniela Renga is a Ph.D. student in Electrical, Electronics and Communications Engineering at the Politecnico di Torino, Italy. She received the Master degree in Computer and Communication Networks Engineering in 2014 from the Politecnico di Torino. In 2004 she received a Laurea degree in Medicine from the Università degli Studi di Torino, Italy. Her research interests are in the field of energy efficient wireless networks, resource management, smart grids and renewable energy sources in the scope of sustainable mobile networks.

Marco Ricca is a research assistant with the Telecommunication Networks Group (TNG) in the Department of Electronics and Telecommunications (DET) of the Politecnico di Torino, Italy. From Politecnico he received the Italian Laurea degree in computer and communication networks engineering in 2008, and a Ph.D. in electrical engineering and communications in 2012. In 2011 Marco was a visiting researcher at Alcatel-Lucent Bell Laboratories in Murray Hill, New Jersey. His primary research interest is in the area of energy-aware control algorithms for packet networks.

Yi Zhang received his B.E. and Ph.D. degree both from Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2007 and 2012, respectively. He was a visiting Ph.D. student at University of California, Davis, U.S.A. from September 2008 to August 2010. He was a postdoctoral researcher at Politecnico di Torino, Italy from November 2012 to May 2015. From June 2015 he has been a postdoctoral researcher at Trinity College Dublin, Ireland. His research interest includes energy-efficient wireless networks and optical networks. He has co-authored over 5 IEEE and Elsevier journal papers. He has served as a Technical Program Committee member for over 10 IEEE and ACM conferences.