Instrumentation for measuring users' goodputs in dense Wi-Fi deployments and capacity-planning rules

José Luis García-Dorado[1], Javier Ramos[1], Francisco J. Gomez-Arribas[1,2], Eduardo Magaña[2,3], and Javier Aracil[1,2]

[1] High Performance Computing and Networking research group, Universidad Autónoma de Madrid, Spain

[2] Naudit High Performance Computing and Networking, S.L., Spain

[3] Telecommunications, Networks and Services Research Group, Universidad Pública de Navarra, Pamplona, Spain

NOTE: This is a version of an unedited manuscript that was accepted for publication. Please, cite as:
José Luis García-Dorado, Javier Ramos, Francisco J. Gomez-Arribas, Eduardo Magaña and Javier Aracil. Instrumentation for measuring users' goodputs in dense Wi-Fi deployments and capacity-planning rules. Wireless Networks, Springer, vol. 26, 2943-2955, (2020).
The final publication is available at:
`https://doi.org/10.1007/s11276-019-02229-7`

Abstract

Before a dense Wi-Fi network is deployed, Wi-Fi providers must be careful with the performance promises they made in their way to win a bidding process. After such deployment takes place, Wi-Fi-network owners—such as public institutions—must verify that the QoS agreements are being fulfilled. We have merged both needs into a low-cost measurement system, a report of measurements at diverse scenarios and a performance prediction tool. The measurement system allows measuring the actual goodput that a set of users are receiving, and it has been used in a number of schools on a national scale. From this experience, we report measurements for different scenarios and diverse factors—which may result of interest to practitioners by themselves. Finally, we translate all the learned lessons to a freely-available capacity-planning tool for forecasting performance given a set of input parameters such as frequency, signal strength and number of users—and so, useful for estimating the cost of future deployments.

Keywords: Wi-Fi performance; goodput; Wi-Fi-network planning; WiFiLytics.

1 Introduction

The current omnipresence of Wi-Fi technology for broadband access in offices, meeting rooms, public transport and homes, to mention only some examples [1], makes it clear its relevance. In this scenario, the lack of deep knowledge of the real Quality of Service (QoS) that Wi-Fi users experience stands out. That is, while typically IT managers focus on coverage—ensuring a certain signal-strength over some reference value or threshold—the effective transmission rates that these users can achieve has

*This work was partially funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund under the project TRÁFICA (MINECO/FEDER TEC2015-69417-C2-1-R) and by Naudit High Performance Computing and Networking under the project ESCUELAS CONECTADAS (Convenios 2018 y 2019 “Análisis de tráfico y supercomputación de sobremesa”, art. 83).

received little attention. However, we believe this is key to ensure the viability of current and future service deployments using Wi-Fi technology [2].

Wi-Fi performance depends mainly on the number of concurrent users, the spatial density of users and the distance/obstacle to the Wireless Access Point (AP) [3]. As intuition suggests, the larger the number of users accessing the same AP is, the larger both the possibility of collisions and the transmission stall times as Wi-Fi typically uses medium access control mechanisms such as Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA) and Request-To-Send (RTS)/Clear-To-Send (CTS). While the former mechanism is based on waiting a random time until a retransmission is sent, the latter is based on an active channel reservation whereby, ideally, only a user is transmitting at a given time while the others remain quiet during such time interval. Similarly, the farther users are placed from an AP, the larger the possibility of coding error due to low signal power. And, finally, a high spatial density of users would foster interference between them. All this leads to larger loss rates and retransmissions, which eventually translates to poor performance [4, 5].

The focus of this paper is to propose a measurement strategy and associated instruments to better understand how such three factors impact on the users' achievable throughput. In more detail, we aim at filling the lack of capacity-planning policies that allow relating the number of users, signal power, and per-AP dispersion to the effective per-user throughput at the application layer—i.e., goodput. We believe that such results may result of interest for both the research community—to achieve a better understanding of Wi-Fi performance—and practitioners that address the task of deploying Wi-Fi networks in schools, universities and other venues where certain QoS levels must be met [6]. Such projects are increasingly common as of today [7] and, however, they are still carried out without having a deep knowledge of what performance Wi-Fi's users are going to achieve, and so making a challenge to plan, in advance, a budget for public tenders. Even more, when SLAs typically establish certain quality levels not only for the average throughput users receive but also for best and, especially, worst cases. That is, not only the aggregate traffic the deployment is capable of carrying but also the quality variance between users, specifically the smallest bandwidth rate that a given user can receive—even more when it is known that TCP users suffer unfairness, especially those with lower signal strengths [8].

Unfortunately, we note that if there is little knowledge about the throughputs Wi-Fi users can empirically achieve on average, even less is known about variance between users' throughputs on the same Wi-Fi network. Among other reasons, we believe that this lack of knowledge is due to the difficulty of performing experimental-driven research with enough number of users/terminals (i.e., high concurrency level) and diversity (i.e., different signal strength conditions and rooms) for significant periods of time, both in terms of measurement procedures and instruments.

Fortunately, we had the opportunity to measure Wi-Fi networks in hundreds of elementary/high schools in Spain during the deployment of Internet connectivity on their buildings as a governmental nationwide project. In particular, we were in charge of gathering throughput measurements in this project, so this is the metric this paper focuses on. Note that throughput has always had an important role in capacity planning, quality assessment and management processes in wireless networks [9, 10].

Our hands-on experience in this project allows us to contribute with the following: First, we describe a low-cost measurement system based on several out-of-the-box commercial APs configured with RTS/CTS access mechanism, 30 domestic-level routers conveniently tuned to behave as users, a traffic generator/sink and a laptop as central manager. As previously introduced, this system was used to assess the technical specifications sheet that telco vendors and public institutions had agreed at a nationwide scale. With it, we report measurements for different numbers of users, frequencies, signal strengths and dispersion in a controlled environment. Then, by elaborating on these measurements, we provide a capacity-planning tool that helps practitioners in the field to design Wi-Fi networks.

2 Low-cost measurement system

While other studies have focused on measuring Wi-Fi performance in a passive way—i.e., real traffic from users but in a non-controllable environment [11]—, or relying on controllable scenarios but using network simulators [12] or analytical approaches [13] instead of real hardware, albeit the dependence of throughput and distance [14], we believe that both characteristics are desirable. To this end, we have developed a system that fulfills both precepts at low-cost for ease of repeatability.

The system comprises four elements: the set of users, the AP that provides Wi-Fi connectivity to such users, a traffic sink/generator and a manager node. Figure 1 illustrates this architecture:

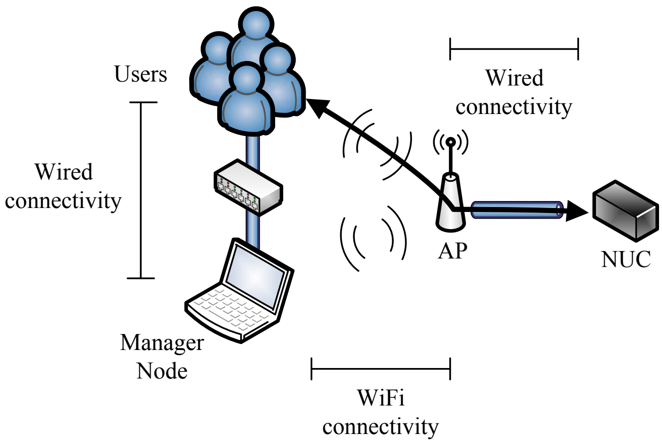


Figure 1: Measurement system architecture

- The traffic sink/generator is an Intel's Next Unit of Computing (NUC) that is a small-form-factor personal PC. We use such a NUC to receive/generate TCP traffic at maximum speed at configurable periods/periodicity/concurrency in both directions exploiting iperf3 [15]. Iperf is a well-tested software to generate standard TCP connections (socket-based) at maximum speed and measure application-layer bandwidth between two nodes in both directions. Finally, the NUC is wire-connected to the AP using three GbE connections.
- The AP is configured with RTS/CTS access mechanism and acts as a bridge forwarding traffic from Wi-Fi interface to Ethernet interfaces and vice versa without performing additional tasks such as NAT or traffic inspection in order to obtain maximum performance. Specifically, we have used APs from six of the most popular vendors.
- The users are a set of 30 Tp-link AC1750 Wireless Dual Band Gigabit Router with 3 external antennas, flashed with dd-wrt [16] GNU/Linux distribution for embedded devices, and distributed evenly on the test rooms under study. The routers feature a Qualcomm QCA9558 chipset for the 2.4 GHz band and a Qualcomm QCA9880-BR4A chipset for the 5 GHz band. Such chipsets present an operational temperature that ranges between 0 and 40^{o} C and an operating humidity range between 10 and 85% which widely covers typical indoor characteristics.
- The manager node serves to configure the parameters of the test (duration, number of users, etc.) and, after tests are finished, it creates records and a webpage showing the throughput—specifically, goodputs—achieved by each user. Also, statistics about the environment and signal level such as Received Signal Strength Indicators (RSSIs), Noise, Modulation and Coding Scheme (MCS) among others are included. By default, the system measures 5-minute goodputs using 12 concurrent connections per user—the next section will show if the latter is a significant figure—and repeats each measurement 3 times.
- We assume that the Wi-Fi physical layer configuration, especially channel bonding, is set to maximize bandwidth, as this is the best setup for providing Wi-Fi coverage in open classrooms. In [17] performance and fairness of IEEE 802.11ac (throughput and jitter) were analyzed with a few greedy clients in a residential and business environment, against Wi-Fi parameters such as channel bonding (20, 40, 80 MHz wide channels), spatial streams, guard intervals and MCSs. In the same mood, the authors in [18] studied the impact of wider channels in network-level performance. As it turns out, wider channels favor Adjacent Channel Interference (ACI), as the number of available channels decreases, which, depending on the traffic scenario, does not pay off for the channel bandwidth increase. This way, we adjust channel bandwidth to 80 MHz in 5 GHz deployments. We note that larger bandwidth channels (160 MHz) are amenable for wireless backhaul communications [19, 20], which is not our use case.

3 Report of measurements

For our measurements, we have distinguished between large and small rooms to account for high and low spatial density—in particular, rooms larger than 15x15 meters are considered low density. In the case of high density, users are close and its signal-strength variance is lower, wherein the low-density case the opposite happens.

With the measurement system and two classrooms of 9 and 300 square meters, we posed two experiments:

- First, we assumed 30 people in each classroom—representing a typical school class, for example in the U.S according to National Center for Education Statistics [21]. In this scenario, we analyze the impact of factors such as the frequency (i.e., 2.4 and 5 GHz), direction (upstream and downstream) and signal-strength power. Considering these factors, we measured the average goodputs achieved per user during a 5-minute experiment. Then, we worked out the aggregate average between all the users as well as the users that fall at percentile 5th and 95th to capture the variance between users—i.e., best and worst users.

- Second, we searched for the distribution of goodput per user according to variations in the number of concurrent users. That is, we started to measure with only a user and, progressively, we added more users.

Namely, given a SLA throughput threshold, the Wi-Fi provider can estimate the number of APs that suffices such SLA, or perhaps as an alternative approach, to have some hints about how much throughput can be delivered for a future project. Further, the second experiment completes the previous figures by showing how throughputs can vary according to the number of users for more/less-populated scenarios. We note that the next section translates both experiments in a prediction tool, which is then evaluated with some extra validation sets.

3.1 Connection concurrency per user

Let us focus, first, on the parameterization of the number of concurrent connections per user for iperf. Essentially, we aim at fully saturating the channel, but initial results showed that a small number of connections (e.g., 1) did not suffice and a far larger number (e.g., 30) may saturate the own CPU capacity of the terminals on the measurement system. As an attempt to find a trade-off, Figure 2 illustrates with some examples the average goodput achieved per user according to the number of concurrent connections from a large set of experiments involving different signal strengths, frequencies, and number of users in both directions. It becomes apparent that the curves tend to show a flat behavior after some simultaneous connections are used.

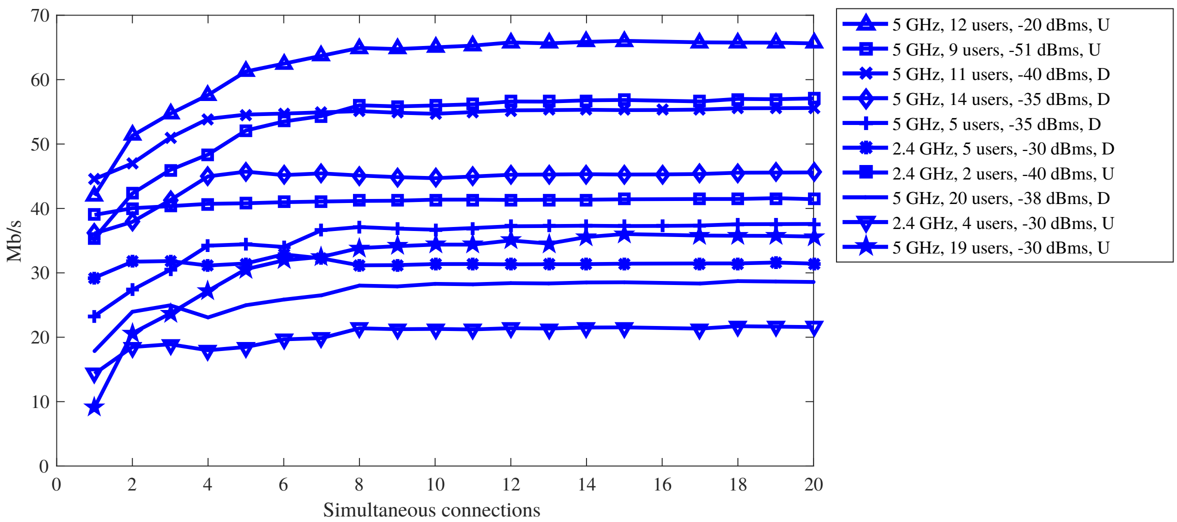


Figure 2: Some illustrative examples of average goodput per user according to the number of simultaneous connections for several signal strengths and number of users in both directions (U = upstream and D = downstream)

In general, less than 10 of such simultaneous connections were not enough to achieve the maximum goodput per user in some of the scenarios. Specifically, scenarios comprising high signal strengths in 5 GHz networks demanded more concurrency, while the same experiments demanded a lower number to saturate users' capacities at 2.4 GHz deployments.

This way, we found 12 as a good trade-off between achieving the maximum user goodputs for all the scenarios and keeping computational load limited. In what follows, 12 concurrent connections per user will be applied in the measurements.

3.2 Signal strength, dispersion and frequencies

Let us detail the results for the first experiment in the 5 GHz deployment. Figure 3(a) depicts average and percentiles 5th/95th as error bars of the goodput each user achieved in this first experiment, in both classrooms and in both directions (at the top, upstream direction—users to sink—, and at the bottom downstream direction—generator to users). In particular, the signal strength per measurement is calculated as the median of the strengths that each user reported—typically, this measurement corresponds to the user located at the centroid of the deployment.

Several observations arise. Average users may expect to achieve goodputs between 5 Mb/s to 25 Mb/s, although better performance is found in the upstream direction of traffic especially at good signal strengths (i.e., over -45 dBms). However, above all, it stands out a larger range between the less and luckier users—in terms of goodput—in the upstream direction than in the counterpart downstream. This range can be as high as 20 Mb/s for the upstream direction, while this figure tends to be smaller than 10 Mb/s in the upstream. Probably, the homogeneity in the download direction has to do with the fact that in this scenario there is only one transmitter and this makes throughput more stable, because of the lesser influence of the contention between multiple transmitters. It is also worth remarking that in the upstream direction there are not significantly better results when the signal strength ranges between -40 to -20 dBms. In this interval, the aggregate goodput peaked at -40 dBms and no better results are achieved for better signal strengths (the next experiment will stress this effect). Finally, by comparing results according to user density, we note that differences are marginal.

Let us now turn our attention to 2.4 GHz deployments whose results are depicted in Figure 3(b). The results show a strong lack of sensitivity of what a user achieves in this deployment on average. This is in the range of 5 Mb/s, being in the downstream direction where some increase of performance is shown for the best signal levels. Regarding, percentiles, it is worth remarking, again, that the ranges for the upstream direction are larger, as users with more throughput may duplicate goodputs with respect to the average user while such users exhibit a 10-times larger goodput than the slowest users. Interestingly, we note that this dispersion may be key to fulfill some requirements that SLAs include.

Finally, by comparing results according to user density, we note that, again, differences are marginal which suggests considering this factor as no relevant.

3.3 Measurements according to the number of users

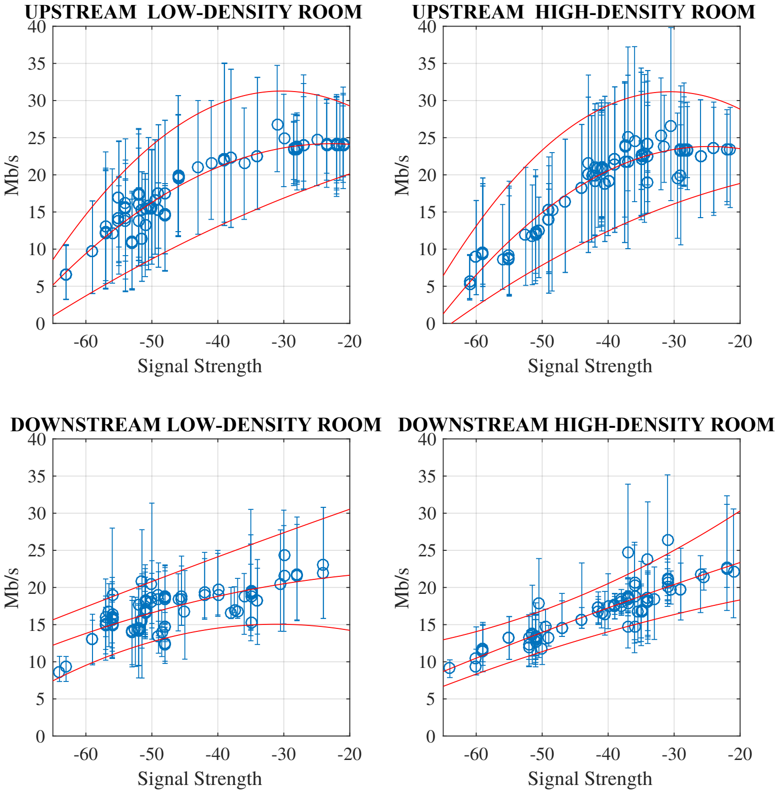
In the second experiment, we vary the number of users between 1 and 30 in different signal-strength environments. Specifically, we have tested in three scenarios, low signal (up to -45 dBms), middle signal (between -45 and -35 dBms) and high signal (signals better than -35 dBms). However, we found that such middle scenario was a simple average of the two others and it does not receive more attention in this section.

Then, Figure 4 shows for high and low signal scenarios (at the top, high-signal scenario) and per direction and density room:

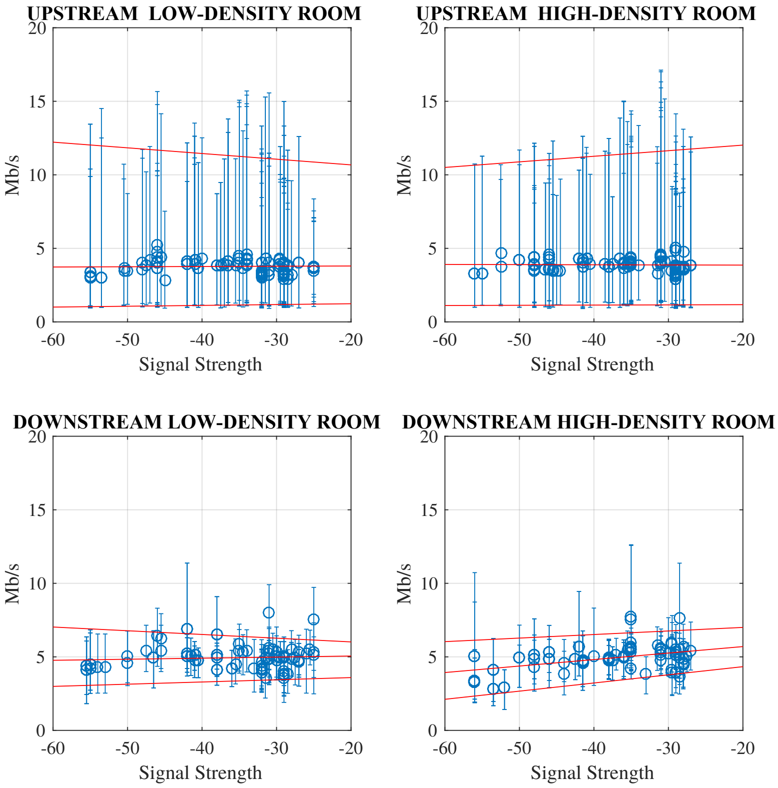
- The total throughput—i.e., the aggregate result of adding goodputs for all the users—for a given number of users (between 1 and 30) achieved. These are the dotted lines at the top of each of the figures, and the right axis represents their values (between 0 and 1 Gb/s).

- For each number of users, the average goodput per user and, as previously, the value for the best and worst users—i.e., percentiles 5th and 95th— are plotted as error bars.

From the former, we see that the traffic aggregate of all users starts with a positive slope up to some 5 users where such aggregate peaks, and starts to decrease with a slight slope—almost flat by the end of the figure. In general, that means that depending on the signal strength there is a limited traffic aggregate that users may share. In the case of the highest signal strengths, such limit was at some 900 Mb/s, probably due to the physical limitation of Wi-Fi protocol such as natural noise and RTS/CTS contention. However, at the low signal-strength scenario, as shown in Figure 4(b), such aggregate was about 400 Mb/s. Moreover, not only aggregate depends on signal strength but also the figure suggests that slopes are different for both scenarios by increasing the number of users.

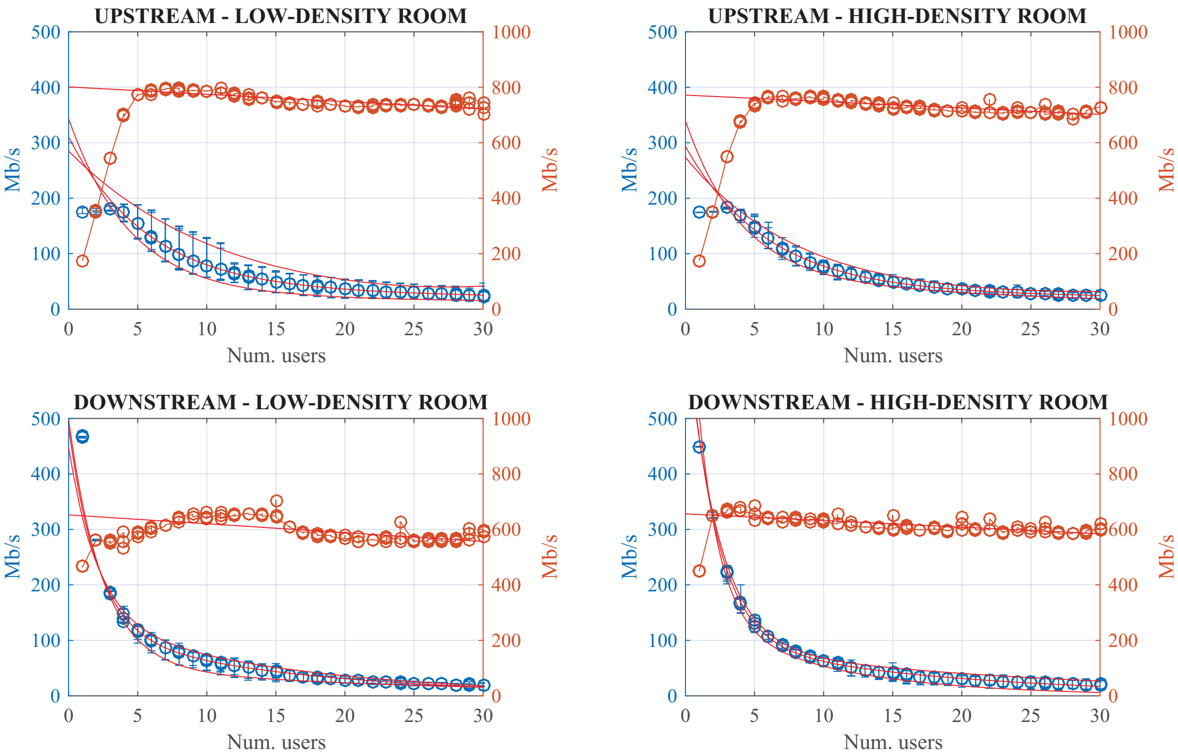


(a) 5 GHz

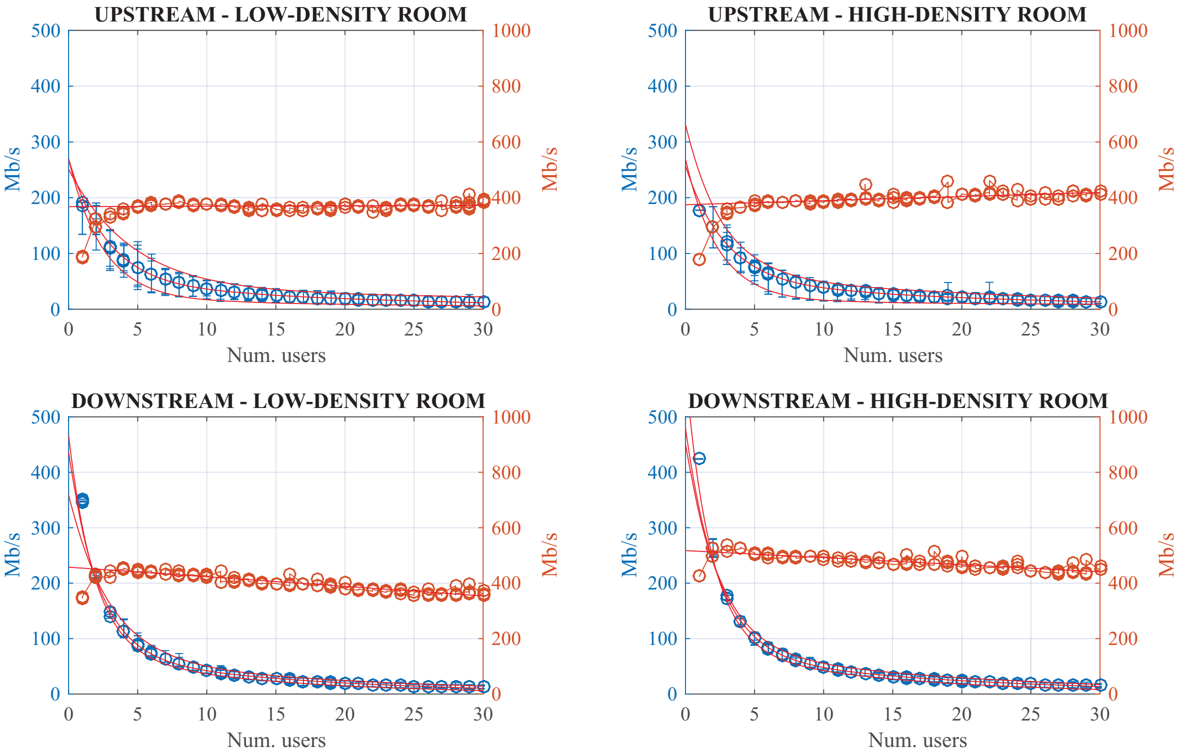


(b) 2.4 GHz

Figure 3: Average and 5th/95th percentiles for goodput per user at 5 GHz (a) and 2.4 GHz (b) deployments for different signal strengths (in dBms) according to the direction (upstream and downstream) and room density (low and high). The continuous lines represent regression curves for the measurements (average and 5th/95th percentiles)



(a) High signal (>-35 dBms)



(b) Low signal (<-45 dBms)

Figure 4: Measured goodputs according to the number of users at the 5 GHz deployment for different signal-strength environments (high and low). The left axis represents goodputs per-user with regression curves for the average and 5th/95th percentiles. The right axis represents aggregate goodput together with its linear regression

Intuitively, on average, traffic shares between users follows a hyperbolic distribution (i.e., each user achieves: total aggregate/number of users), however we have to add the fact that such aggregate tends to decrease slightly with the number of users and that the behavior is different with few users. That is, by adding more users, such aggregate does not increase or simply keeps constant—in this case, each user receiving a smaller fraction—but the interference between users themselves makes aggregate decrease—and so, less amount of bandwidth to share between more users. Besides, the result for a low number of users shows that a few of them are incapable of saturating the transmission channel—i.e., a limit of some few hundreds of Mb/s per user in our measurements.

It is worth remarking that this latter point—the throughput that a unique user may generate/receive—may depend on the own quality of the terminals' antennas. Actually, we carried out tests with Mac laptops that showed a limit per user over such figures. Other terminals such as cell phones that typically use 2x2 mode (2 antennas using 2 spatial streams) present per-user throughput far below our terminals—which operate in 3x3 mode. However, we note that, even with high-end laptops or using dedicated external antennas, a single user is not able to fully saturate the channel. In any case, note that the key message of this study is not the goodput that a single user or pair of them may achieve as it depends on antennas and terminal characteristics. Instead, we focus on Wi-Fi providers—for instance, telcos—that deploy dense Wi-Fi networks for schools, transport hubs, etc. with more than this amount of users. In conclusion, goodputs with a few users are difficult to generalize, but over such a figure of three or four users, things become more homogeneous.

We now pay attention to the left axis of the figure, it shows the results that each user, individually, achieved as well as representation for the best and worst users. In other words, we focus on how homogeneously the traffic aggregate is divided. We remark that this probably represents a significant metric for most of the SLAs. Furthermore, the figure shows again that the upstream direction behaves more heterogeneously with some users account for up to 70% more than others—which was relatively expected given that current medium access protocol mechanisms cannot provide fairness [22]. Although, when the number of users increases this dispersion is becoming less significant.

We note that due to the scales, it may result difficult to appreciate the specific values of each measurement in the previous figures, fortunately, the next section will introduce WiFiLytics. WiFiLytics is a tool that exploits this rich set of measurements to predict throughputs per user according to frequency, number of users and signal strength. WiFiLytics is available at [23], this way the specific values of any point in the figures can be easily retrieved. For example, assuming upstream direction in a 5 GHz deployment with high signal, by decreasing the number of users from 30 to 15 users goodputs increase in the average per user about a 40%, for 5th percentile more than 50% and roughly 20% at percentile 95. This emphasizes the relevance of paying attention to percentiles as behavior is far from being homogeneous in the worst, best and average cases.

Finally, Figure 5 shows the same measure but for a 2.4 GHz deployment for the high-signal strength scenario. In general, conclusions are the same. Although the traffic aggregate behaves flatter with the number of users and, as expected, the total aggregate is lower—between some 100-150 Mb/s. We note that WiFiLytics also includes results for 2.4 GHz deployments for low/middle signal strengths for a more detailed inspection of data.

3.4 Learned lessons

Finally, let us summarize the main lessons learned after the review of the measurements shown previously. We believe can result of interest for both owners of Wi-Fi-network deployments and practitioners in the area:

- We have found very significant differences between the most and the least lucky users. Therefore, when verifying a Wi-Fi deployment SLA, not only average aggregate goodput is relevant but also the worst clients' goodput must be taken into account as a quality indicator.

- By measuring performance at different rooms we paid attention to the users' density however, we obtained equivalent results in the high and low user-density rooms used in the tests.

- The total traffic aggregate decreases with the number of users and so, less amount of bandwidth to share between more users

- More than a user is needed to saturate the channel capacity depending on the quality of devices' antennas.

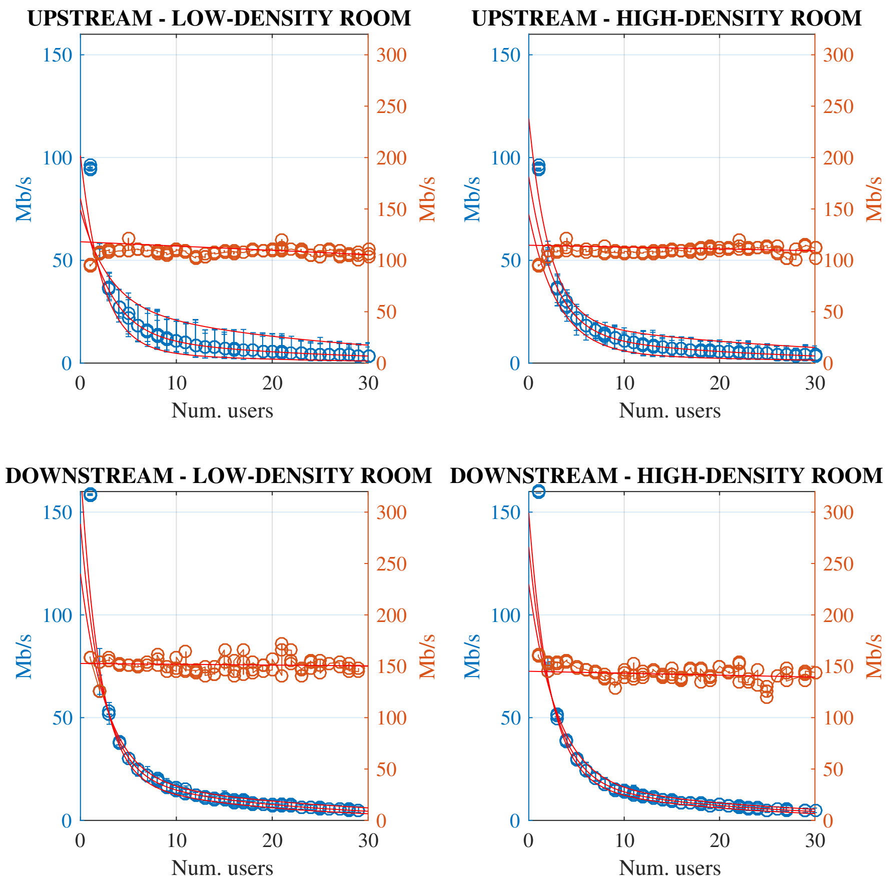


Figure 5: Measured goodputs according to the number of users at the 2.4 GHz deployment at the high-signal strength scenario. The left axis represents goodputs per-user with regression curves for the average and 5th/95th percentiles. The right axis represents aggregate goodput together with its linear regression

- As expected, 5 GHz deployments are extremely sensitive to the signal strength, while the same does not apply to 2.4 deployments.
- Performance differs in upstream and downstream directions, due to the lesser influence of the contention between multiple transmitters.
- Given a set of parameters such as frequency, signal strength and number of users, the measurements have shown a consistent behavior, which calls for its modeling and, subsequent, forecasting.

This way, in the next section, we translate all these learned lessons and the rich set of gathered measurements to a freely-available tool for forecasting performance.

4 Capacity-planning tool

Paying attention to the last three figures, we note that all of them, in addition to the measurements we have already reviewed, depict regression models as continuous lines. In more detail:

- Figure 3 illustrates the modeling, for each trio of direction, density and frequency, of the goodput on average, 5th and 95th percentiles according to the measured signal strength—in particular for 30 users. For the 5 GHz measurements, the analytical model that has shown the best goodness-of-fit is a quadratic function:

  $$\text{Goodput}(x = \text{signal-strength}) = p1 \cdot x^2 + p2 \cdot x + p3, \tag{1}$$

  where $p1$, $p2$ and $p3$ are parameters that we estimate empirically from measurements.

  However, the same does not apply to the 2.4 GHz deployment. Actually, from previous results, we have already known about the insensibility of performance to the signal-strengths. In this case, the fittest models are simple linear regressions:

  $$\text{Goodput}(x = \text{signal-strength}) = p1 \cdot x + p2, \tag{2}$$

  which describes a straight line with slope $p1$ and constant term $p2$.

  Remarkably, the confidence intervals for slopes included zero, this means that the correlation between the two variables, signal strengths and goodput, is not statistically relevant in the ranges we are studying.

- The figures 4 and 5 illustrate the modeling for goodputs according to the number of users. These regressions have been modeled by two-term exponential models for both 2.4 and 5 GHz scenarios:

  $$\text{Goodput}(x = \text{number-of-users}) = p1 \cdot e^{(p2 \cdot x)} + p3 \cdot e^{(p4 \cdot x)} + p5, \tag{3}$$

  where $p1$ ... $p5$ are parameters empirically estimated from measurements.

  Note that these two exponential terms account for both tendencies reviewed in the previous section. When the number of users is high—and traffic aggregate almost constant—, a hyperbolic function that distributes the traffic aggregate between the active users, and when the number of users is lower, another capturing that the traffic aggregate is increasing/decreasing. In between these two states, a proportion of both terms is applied. Additionally, note that the figures include a linear regression for traffic aggregates (i.e., the total exchanged traffic by all users), whereby negative slopes become evident.

With all the previous results in mind, let us define design guidelines for a capacity-planning tool for Wi-Fi deployments.

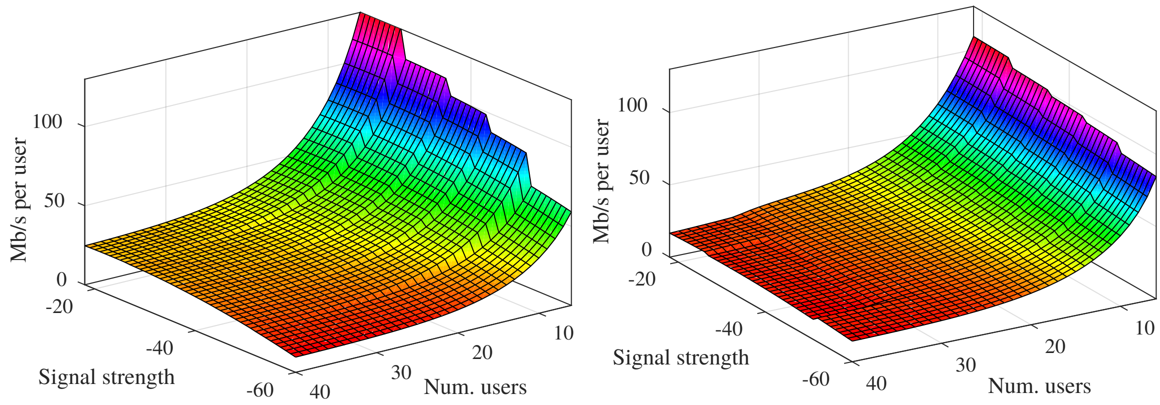
4.1 WiFiLytics

Once modeled and estimated the parameters for the distribution of user goodputs on factors such as frequency, direction and density according to both the signal strength and number of users using the reported set of measurements, we propose to merge them in a prediction tool for permanence in Wi-Fi environments.

Note that the first regression (signal strengths) allows us to estimate the performance for a fixed set of 30 users, and the second regression (number of users) allows us to estimate the fraction of goodput gained/lost when varying the number of users. This way, once modeled and estimated the parameters using the reported set of measurements, we propose to merge them in a prediction tool for permanence in Wi-Fi environments, to which we have named as WiFiLytics (freely available at [23]).

Therefore, given a frequency, signal strength, spatial density and number of users, WiFiLytics calculates the goodput for 30 users—the default number of users—for the given signal strength. Then, it increases/decreases such goodput according to the increase/decrease that the regression for the number of users indicates from 30 users to the specific number of users requested—applying regression according to groups of high, middle and low signal strength. And all this for the average, 5th and 95th percentiles estimates.

As an example to illustrate the estimates that WiFiLytics may provide, Figure 6 shows the average goodput a user may expect to achieve according to the number of other users and the signal quality in a 5 GHz network.



(a) Upstream (b) Downstream

Figure 6: Average per-user goodput estimates according to the number of users and signals strengths (in dBms) at a 5 GHz scenario assuming low-density of users

4.2 Evaluation

As an evaluation, we focus on both the goodness-of-fit of regressions and the accuracy of applying WiFiLytics' models in a new set of measurements—i.e., validation set—in terms of the coefficient of determination (R^2) between measured values and model and mean percentage error (MPE) between estimates and real measurements, respectively.

Regarding regressions, tables 1 and 2 show the R^2 goodness-of-fit coefficients for each regression (middle power-strength scenario is not shown separately as their results are equivalent to the others in practical terms) for both frequencies, respectively. Several observations arise:

- The regression for the number of users shows significant fittings for both frequencies and all percentiles ranging between 0.999 and 0.93.
- With respect to the regression for the signal strength:
  - The results for 5 GHz are significant for upstream direction ($0.92 - 0.78$) and fairly significant for downstream ($0.73 - 0.63$).
  - On the other hand, in 2.4 GHz the R^2 coefficients were non-significant. This means that the linear regressions do not add, almost, information to the estimate beyond the simple

mean. This was expected given the results that Figure 3(b) showed, this, simply, makes the prediction tool simpler in this aspect.

Let us, finally, assess how these models translate in numerically estimates and how far are from measurements not used as training for the modeling. We have carried out a set of new measurements picking, randomly, between the two frequencies and placing APs in different rooms, distance from users and positions—hence varying signal strengths. The mean error for WiFiLytics fairly ranged between 4% and 12% for best and worst cases respectively as shown in the rows of Table 3. In 5 GHz scenarios, these figures translate in absolute errors of a few Mb/s and, in the case of 2.4 GHz scenario, even less. Consequently, this emphasizes WiFiLytics' usefulness as a guide for future Wi-Fi deployments.

Table 1: Adjusted R^2 goodness-of-fit coefficient for regressions for average, 5th and 95th percentiles on signal strengths and number of users measurements for the 5GHz deployment

			5 GHz					
			Upstream			Downstream		
			5th	Average	95th	5th	Average	95th
Regression	Signal strength		0.7796	0.9235	0.8646	0.6303	0.7339	0.6683
	Number of users	High power-strength	0.9962	0.9989	0.9885	0.9963	0.9938	0.9908
		Low power-strength	0.9669	0.9980	0.9760	0.9968	0.9986	0.9944

Table 2: Adjusted R^2 goodness-of-fit coefficient for regressions for average, 5th and 95th percentiles on signal strengths and number of users measurements for the 2.4GHz deployment

			2.4 GHz					
			Upstream			Downstream		
			5th	Average	95th	5th	Average	95th
Regression	Signal strength		0.0385	0.0084	0.0120	0.2105	0.0952	0.0190
	Number of users	High power-strength	0.9838	0.9973	0.9775	0.9974	0.9977	0.9941
		Low power-strength	0.9835	0.9907	0.9506	0.9866	0.9878	0.9339

Table 3: Mean percentage error on validation set for both 2.4 and 5 GHz deployments

		Upstream			Downstream		
		5th	Average	95th	5th	Average	95th
Mean Percentage Error	5 GHz	7.1	6.7	8.9	8.7	8.9	10.1
	2.4 GHz	12.3	7.4	10.9	9.7	4.4	9.9

5 Related work

Finally, let us put in context the contributions and results shown throughout this work.

In [24] the authors study the performance of TCP over a Wi-Fi with RTS/CTS activated. They consider download and upload TCP long-lived connections and conclude that as the AP buffer decreases the download throughput decreases and the upload throughput increases even in the presence of wireless channel error. An analysis of TCP download performance in dense scenarios is performed in [11], through experimentally-driven analysis and simulation with a large number of Wi-Fi clients. The authors conclude that the upload traffic, even if composed by ACKs only, has a large influence in the download throughput and propose to orchestrate downloads in round-robin. In [8], it was reported that TCP clients suffer unfairness, especially those with lower signal strengths.

The above works provide valuable insight into the dynamics of the TCP protocol in Wi-Fi networks but they lack a capacity-planning rule. Furthermore, the Wi-Fi version evaluated is not the current IEEE 802.1ac standard at 5 GHz.

As for measurement instruments, we distinguish between client emulators and those that feature separate physical equipment (such as a laptop) per client. In this first group, there are several commercial examples from Ixia [25] and Candela [26]. The clients are virtualized into a single physical appliance, which features several Wi-Fi antennas. If the number of clients is larger than the number of antennas, some degree of serialization necessarily happens at the physical layer because clients cannot transmit concurrently. Furthermore, all the emulated clients receive the same Wi-Fi signal, namely with the same power level, which is not the case in a large classroom with physically dispersed clients. We note that the advantage of emulation is that the equipment is less bulky and can be easily deployed and ported, but it provides an emulation of the client dynamics, which is not as accurate as having physical clients deployed for the measurement test.

In [14], an experimental measurement study of indoor Wi-Fi in a floor plant is performed, with special emphasis on the distance from the client to the AP, for IEEE 802.11n and 802.11ac. Tests are performed with five clients and using UDP as a transport protocol. The dependence of throughput with distance is rather striking, which reinforces the point that using physical clients instead of emulated ones provides more realistic results.

In [17] performance and fairness of IEEE 802.11ac (application-level throughput and jitter) are analyzed with a few greedy clients in a residential and business environment, against Wi-Fi parameters such as channel bonding (20, 40, 80 MHz wide channels), spatial streams, guard intervals and modulation and coding schemes. In the same mood, the authors in [18] study the impact of wider channels in network-level performance. As it turns out, wider channels favor adjacent channel interference (ACI), as the number of available channels decreases, which, depending on the traffic scenario, does not pay off for the channel bandwidth increase.

In [12], the authors perform a TCP throughput analysis varying the number of TCP clients using the Ix-Chariot tools from IXIA, with clients possibly virtualized into several laptops, but not with a single laptop per client, then avoiding full concurrency in the physical layer. In [22] a testbed of 20 clients and 6 APs is presented that serves to demonstrate that fairness cannot be achieved with standard MAC mechanisms, as the throughput received largely depends on the heterogeneity in selecting channel access parameters among neighboring wireless access points. As a result, the authors propose a fair medium-access protocol.

As shown, the experimentally driven analysis currently available in the state of the art does not provide capacity-planning rules for Wi-Fi dimensioning at the application layer, as we contribute in this paper.

6 Conclusions

Throughout this article, we have shared our experience in the task of designing measurement instruments and measuring Wi-Fi network throughput in a number of public schools. We believe that this experience can result of interest for both owners of Wi-Fi-network deployments and practitioners in the area.

That is, we have first presented a guide on how to deploy Wi-Fi measurement instruments at low cost. This step is key for institutions that have signed a SLA with Wi-Fi providers and are willing to assess that such SLA is being fulfilled. Likewise, Wi-Fi providers may find of interest such a system to evaluate the performance of their own deployments and so avoiding future penalties.

Then, we have changed our focus towards Wi-Fi providers' viewpoint. They faced the opposite problem on its target of winning a bidding process: how much quality they can offer given a fixed budget or, given some fixed specifications of quality how the cost of the deployment would be. For them, we have reported a number of real measurements observing factors such as the number of users, signal strength, density among others as well as made public a capacity-planning tool for forecasting performance, WiFiLytics. This way, Wi-Fi providers may have beforehand some hints of both the number of APs and their locations they would need to achieve the requested quality and, also, the cost over which to give a profitable proposal.

Finally, the measurements mentioned here as well as the final tool WiFiLytics are freely available at [23].

References

[1] J. Witzel. 802.11, it's out there for free... if you know where to look. *IEEE Instrumentation Measurement Magazine*, 10(5):54–57, 2007.

[2] Paul S. Henry and Hui Luo. WiFi: what's next? *IEEE Communications Magazine*, 40(12):66–72, 2002.

[3] Xiaoling Yang and Bing Chen. A novel method for measurement points selection in access points localization. *Wireless Networks*, 24(1):257–270, 2018.

[4] Amit P Jardosh, Krishna N Ramachandran, Kevin C Almeroth, and Elizabeth M Belding-Royer. Understanding congestion in IEEE 802.11b wireless networks. In *ACM SIGCOMM Conference on Internet Measurement*, pages 25–25, 2005.

[5] L. Angrisani and C. Narduzzi. Testing communication and computer networks: an overview. *IEEE Instrumentation Measurement Magazine*, 11(5):12–24, 2008.

[6] Vladimir Brik, Shravan Rayanchu, Sharad Saha, Sayandeep Sen, Vivek Shrivastava, and Suman Banerjee. A measurement study of a commercial-grade urban WiFi mesh. In *ACM SIGCOMM Conference on Internet Measurement*, pages 111–124, 2008.

[7] Bhaskaran Raman and Kameswari Chebrolu. Experiences in using WiFi for rural Internet in India. *IEEE Communications Magazine*, 45(1):104–110, 2007.

[8] Vasileios P. Kemerlis, Eleftherios C. Stefanis, George Xylomenos, and George C. Polyzos. Throughput unfairness in TCP over WiFi. In *Conference on Wireless On-demand Network Systems and Services*, pages 26–31, 2006.

[9] Dimitrios Koutsonikolas and Y. Charlie Hu. On the feasibility of bandwidth estimation in wireless access networks. *Wireless Networks*, 17(6):1561–1580, 2011.

[10] Saravanan Kandasamy, Ricardo Morla, Patrícia Ramos, and Manuel Ricardo. Predicting throughput in IEEE 802.11 based wireless networks using directional antenna. *Wireless Networks*, 25(4):1567–1584, 2019.

[11] Mukulika Maity, Bhaskaran Raman, and Mythili Vutukuru. TCP download performance in dense WiFi scenarios: Analysis and solution. *IEEE Transactions on Mobile Computing*, 16(1):213–227, 2017.

[12] Sneha V. Sangolli and Thyagarajan Jayavignesh. TCP throughput measurement and comparison of IEEE 802.11 legacy, IEEE 802.11n and IEEE 802.11ac standards. *Indian Journal of Science and Technology*, 8(20), 2015.

[13] Giuseppe Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, 2000.

[14] Mihaela-Diana Dianu, Janne Riihijarvi, and Marina Petrova. Measurement-based study of the performance of IEEE 802.11ac in an indoor environment. In *IEEE International Conference on Communications*, pages 5771–5776, 2014.

[15] Ajay Tirumala, Mark Gates, Feng Qin, Jon Dugan, and Jim Ferguson. Iperf - the TCP/UDP bandwidth measurement tool. `https://github.com/esnet/iperf`, 2019.

[16] DD-WRT. Opensource firmware for wlan routers and embedded systems. `https://www.dd-wrt.com`, 2019.

[17] Lito Kriara, Edgar Costa Molero, and Thomas R. Gross. Evaluating 802.11ac features in indoor WLAN: An empirical study of performance and fairness. In *ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation, and Characterization*, pages 17–24, 2016.

[18] Ljiljana Simic, Janne Riihijärvi, and Petri Mähönen. Measurement study of IEEE 802.11ac Wi-Fi performance in high density indoor deployments: Are wider channels always better? In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 1–9, 2017.

[19] Ruizhi Liao, Boris Bellalta, Jaume Barcelo, Victor Valls, and Miquel Oliver. Performance analysis of IEEE 802.11ac wireless backhaul networks in saturated conditions. *EURASIP Journal on Wireless Communications and Networking*, 2013(226), 2013.

[20] Raja Karmakar, Samiran Chattopadhyay, and Sandip Chakraborty. Impact of IEEE 802.11n/ac PHY/MAC High Throughput Enhancements on Transport and Application Protocols—A Survey. *IEEE Communications Surveys Tutorials*, 19(4):2050–2091, 2017.

[21] National Center for Education Statistics. Average class size in public primary schools, middle schools, high schools, and schools. `https://nces.ed.gov/surveys/sass/tables/sass1112_2013314_t1s_007.asp` [19 Dec. 2019].

[22] Raja Karmakar, Samiran Chattopadhyay, and Sandip Chakraborty. Channel access fairness in IEEE 802.11ac: A retrospective analysis and protocol enhancement. In *ACM International Symposium on Mobility Management and Wireless Access*, pages 51–58, 2016.

[23] HPCN-UAM. WiFiLytics. `https://github.com/jlgarciadorado/wifilytics`, 2019.

[24] Shiva Raj Pokhrel, Manoj Panda, Hai L. Vu, and Michel Mandjes. TCP performance over WiFi: joint impact of buffer and channel losses. *IEEE Transactions on Mobile Computing*, 15(5):1279–1291, 2016.

[25] Ixia. Ixia blackbook: 802.11ac Wi-Fi benchmarking. `http://www.ixiacom.com/blackbook`, 2017.

[26] Candela Tech. Testing WiFi networks with LANforge. `https://www.candelatech.com/lf_wifi.php`, 2019.